Document Number – IN0092-109(I)
CI Number – SS0051A
Date: 16 Aug 2001

ALGORITHM THEORETICAL BASIS DOCUMENT (ATBD)
(IR TOTAL COLUMN OZONE)

for the

Ozone Mapping and Profiler Suite (OMPS)
of the
National Polar-Orbiting Operational Environmental
Satellite System (NPOESS) Program

Contract Number: F04701-99-C-0044
CDRL A012

PREPARED FOR:
NPOESS Program Office

Prepared by
Ball Aerospace & Technologies Corp. (BATC)
Boulder, CO 80306
CAGE Code 13993

PREPARED BY
Hilary E. Snell

_____
OMPS AER Project Manager

APPROVED:
Juan Rodriguez

_____
OMPS Systems Engineer

APPROVED:
Roger Scarlotti

_____
OMPS Program Manager

Project Release _____

Revision Record

Revision	Date	Reasons for Revision	Project Released Date
Draft	March 1, 1999	1st Draft Release	
Revised Draft	April 6, 1999	Proposal Improvements	
Initial Release	January 2000	Updates as of RRR1	
Version 2	October 2000	Updates as of RRR2	
Rev A-Version 3	Aug 16, 2001	Updates as of RRR3	

Table of Contents

List of Figures

List of Tables

1 Introduction

The Ozone Mapping and Profiler Suite (OMPS) is an important component of the National Polar-Orbiting operational Environmental Satellite System (NPOESS). The OMPS mission is to provide the NPOESS users with data products that describe the vertical, horizontal and temporal distribution of ozone in the Earth's atmosphere. These data (Environmental Data Records, or EDRs) are derived from the space-borne ultraviolet and visible observations of a two-sensor system. To enhance the global coverage and provide a means of cross-validation, the total column ozone EDR is also derived from measurements made by the NPOESS-CrIS (Cross-Track Infrared Sounder).

1.1 Scope

An individual document has been developed for each of the four OMPS algorithms. The OMPS algorithms include the following:

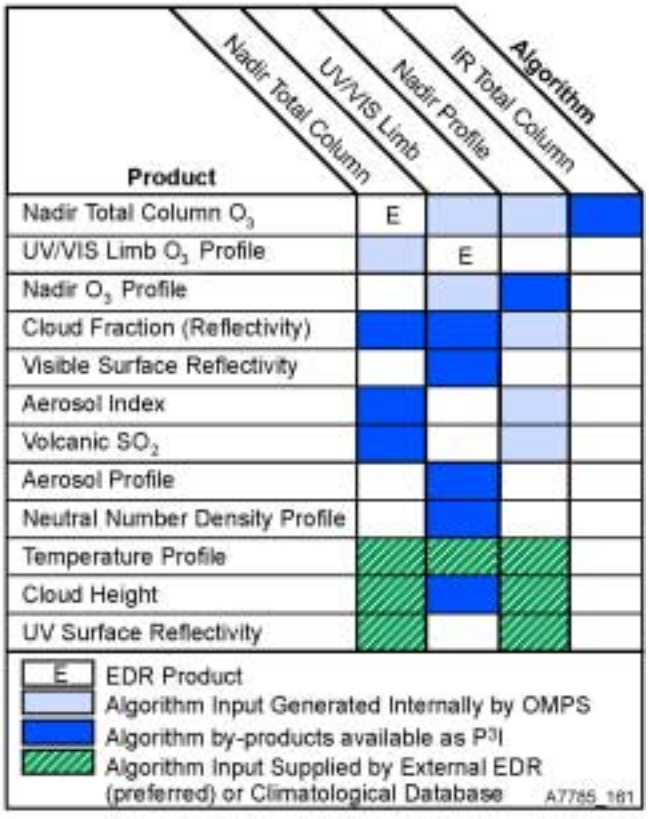
1 **The UV Nadir Total Column Ozone Algorithm** is adapted from the heritage TOMS version 7 algorithm. We have included modular enhancements to meet EDR requirements and to provide for graceful degradation.

2 **The UV Nadir Profile Ozone Algorithm** is adopted from the heritage SBUV/2 operational algorithm. The ozone profile from this algorithm not only provides an initialization for the UV/VIS Limb Profile Algorithm but also provides a link to the heritage twenty-year ozone profile data set.

3 **The UV/VIS Limb Profile Ozone Algorithm** is adapted from the heritage SOLSE/LORE algorithm. We have included modular enhancements to achieve EDR requirements and to provide graceful degradation.

4 **The IR Total Column Ozone Algorithm** is adapted from heritage algorithms used for TOVS, CIRRIS-1A, and EOS-TES data. In order to improve the performance of the ozone retrieval, auxiliary parameters such as temperature and moisture profiles, surface emissivity, and surface skin temperature are retrieved simultaneously with the ozone column amount. The IR ozone values are reported at locations that complement the UV nadir total ozone values (i.e., for SZA greater than 80 degrees and within the SAA).

Table 1.1-1. The OMPS algorithms take advantage of internally generated products while minimizing dependence on external data



These algorithms and their output

products are summarized in **Table 1.1-1**.

1.2 System Description and Objectives

The OMPS consists of two sensors designed to measure the total column ozone and ozone profiles. The sensors operate at UV/Visible/NIR wavelengths and provide measurements for solar zenith angles less than 80 degrees. The IR total column ozone algorithm (IR-TCA) will be used with data from the Cross-Track Infrared Sounder (CrIS) to enhance the global coverage of the OMPS by providing accurate ozone retrievals for solar zenith angles greater than 80 degrees (i.e. nighttime). In addition to enhancing the global coverage of the OMPS, daytime use of the IR-TCA will produce an independent estimate of the total column ozone amount that can be used for the cross-calibration and validation with the measurements made by the UV-nadir sensor. This is particularly useful within the South Atlantic Anomaly (SAA). Further, this algorithm will provide a robust backup of this key parameter in the event of a failure of the UV-nadir sensor.

The primary spectral region of interest for ozone in the IR is 1000 – 1065 cm^{-1}. However, this region is also influenced by the water vapor continuum, which is highly sensitive to temperature [Clough, *et al.*, 1995]. Further, this region is not entirely opaque to a space-borne sensor, and knowledge of surface parameters (skin temperature and emissivity) is important for the radiance calculation. Consequently, the IR-TCA relies on the accurate knowledge of the atmospheric state (temperature, water vapor, surface parameters, and cloud properties) to enhance the retrieval of ozone.

The IR-TCA ingests both CrIS sensor data records (SDRs) and environmental data records (EDRs). The EDRs (pressure, temperature and water vapor profiles) are used as part of the first-guess profile inputs to the forward model. The first-guess parameters for the ozone profile and surface parameters are obtained from a climatological database. The algorithm requires no other external data. Using the CrIS EDRs as a first guess provides an accurate estimate of the temperature and water vapor profiles, and enhances the speed and accuracy of the IR-TCA. Currently the algorithm uses three bands of the CrIS instrument (see Section 3). Once a CrIS design is finalized, a channel selection may be performed to determine the optimum number of CrIS channels to use in the retrieval. In a graceful degradation mode, only SDR data from Band 1 of CrIS (nominally 650 – 1095 cm^{-1}) is required.

The forward radiance model uses the Optimal Spectral Sampling (OSS) algorithm. The OSS is a highly modular, monochromatic radiative transfer model that can be tuned to optimize both computational speed and accuracy. The version of this algorithm for the IR-TCA has been validated with AER's line-by-line radiative transfer model (LBLRTM), and the calculations agree to within 0.05 K in brightness temperature. LBLRTM is a state-of-the-art model that has been extensively validated against atmospheric measurements [Clough, et al., 1995].

The radiance inversion module of the IR-TCA is based upon the optimal estimation (OE) technique [Rodgers, 1976; Rodgers, 1990]. This is a robust and proven technique that allows accurate accounting of both sensor and atmospheric characteristics inherent in the measurement by incorporating both the physical radiative properties of the measurement as well as known statistical correlations between the atmospheric parameters. These statistical correlations can be updated as necessary through the use of EDRs from the OMPS and other NPOESS sensors.

Another key feature of the IR-TCA is the advanced treatment of clouds. The algorithm is modular and has been designed and constructed such that cloud parameters can be easily incorporated into the retrieval procedure. Once a design has been selected for the CrIS sensor, simulations will be performed to understand the degree to which clouds will impact the measured radiances and, more importantly, the retrieval of ozone. We envision that the final product will be an algorithm that is sufficiently robust to accurately treat homogeneous overcast, partial cloudiness, and complex fields with optically thin cirrus clouds.

The objective of this Algorithm Theoretical Basis Document (ATBD) is to describe the algorithm used to retrieve the OMPS Total Column Ozone Product using EDRs and SDRs from the CrIS Sensor. This product consists primarily of the total ozone in a column of air from 0 to 60 km and observed for all solar zenith angle viewing conditions greater than or equal to 80 degrees. This document:

a) identifies the sources of input data that are required by the algorithm, including assumptions about the CrIS sensor and its SDRs and EDRs;
b) provides the physical theory and mathematical background underlying the use of this information;
c) describes practical considerations affecting the detailed algorithm development;
d) lists the assumptions employed in the algorithm retrieval process, and describes the EDR products and additional algorithm by-products;
e) details expected sensor and algorithm errors (accuracy and precision);
f) discusses the use of calibration datasets; and
g) outlines our test and validation approaches.

1.3 Data Products

The IR total column ozone algorithm (IR-TCA) will provide a measure of the total column ozone amount from the surface to 60 km. The result will be given in Dobson Units (DU). Because the CrIS sensor has ~15 km spatial resolution, a (TBS) mapping algorithm will be used to convert the retrieved ozone amounts from the CrIS spatial grid to the 50 km grid required for OMPS.

1.4 Applicable Documents

1.4.1 Controlling Documents

1. OMPS System Specification – Document Number 542798
2. OMPS Algorithm Development Specification – Document Number 542808

1.4.2 OMPS Reference Documents

1. OMPS Algorithm Theoretical Basis Document: Nadir Total Column Ozone Algorithm—Document Number IN0092A-106.
2. OMPS Algorithm Theoretical Basis Document: UV/VIS Limb Profile Ozone Algorithm—Document Number IN0092A-107.
3. OMPS Algorithm Theoretical Basis Document: Nadir Profile Ozone Algorithm—Document Number IN0092A-108.
4. OMPS Algorithm Theoretical Basis Document: IR Total Column Ozone Algorithm—Document Number IN0092A-109.

1.5 Revision History

The original version of the IR ATBD was dated March 1, 1999. ***The ATBD is intended to be a working document that will not be finalized until just prior to the critical design review (CDR), although many of the sections may remain unchanged. Consequently some sections will be sparse or incomplete until CDR.*** Revisions to this document will be made as new information about the CrIS sensor or the algorithms becomes available, and as performance information is verified and validated. Significant changes to content or format for a given version of this document are given in **Table 1.5-1**.

Table 1.5-1. Revision History of This Document

Revision	Release Date	Section	Change Type
Phase 2, Version 1.0	January 2000	1.5	Added text and table
		2.2	Updated CrIS assumptions
		2.3.1.2	Clarifications; fixed equation 7; added solar radiance section
		3.1 – 3.2 and 5.1	All assumptions about SDRs and EDRs now given in Section 2.2; specific requirements for the CrIS are given in Section 5.
		7.0	Clarification about current results.
		8.1 - 8.3	Updates and clarifications.
		8.6	New section on algorithm timing.
		B.3	Disclaimer on code description
Phase 2, Version 2.0	October 2000	7.0	Re-organized and re-written to reflect current accuracy and precision estimates
		2.2	Minor updates based on information provided by ITT about the CrIS sensor
		2.3.1.2.3	Updated to reflect addition of solar source function
Phase 2, Version 3.0 (AER doc #: P869-ATBD-I-20010621)	June 2001	All	Consistent numbering of figures and tables; hyper-linking of references to figures, tables, and section numbers
		2.1	Replaced existing figures with new figures that better explain the retrieval problem
		2.4	New section regarding clouds
		2.6	New section regarding retrieval algorithm's use of principal components analysis (PCA)
		3.2.3	Clarification about channels used by algorithm
		7.0	Reorganized to provide easier access to information; updated baseline performance results
		8.0	Additional text; moved timing tests to appendix
		Appendices	New sections for: quality control; merging of CrIS and OMPS-IR EDR algorithms; test profile data file format; timing tests
		Science Code	Detailed description moved from the ATBD into a separate file

Note: Table 1.5-1 contains major changes only; corrections/updates of equation numbers, changes in formatting, and the addition/updating of reference material is not included in this list. Also, "track changes" has been used to provide differences between this version and the previous version. However, only substantive changes are tracked (i.e. the movement of sections will not show up as a change, except where changes have been made within a section). ***For certain situations the use of this MSWord feature results in non-sequential numbering of figures and tables. This problem will be fixed for the final version.***

1.6 Contributing Authors

Contributors to each of the four OMPS ATBDs include:

Lead:	Nadir Total Column Ozone Algorithm	Colin Seftor (Raytheon)
Lead:	Nadir Profile Ozone Algorithm	Charles Wellemeyer (Raytheon)
Lead:	UV/VIS Limb Profile Ozone Algorithm	Jack Larsen (Raytheon)
Lead:	IR Total Column Ozone Algorithm	Hilary Snell (AER)

Other contributors include:

Susan Beresford	AER, Inc.
Brent Canova	Ball Aerospace & Technologies Corp.
Kelly Chance	Harvard-Smithsonian
David Flittner	University of Arizona
Jennifer Hegarty	AER, Inc.
Benjamin Herman	University of Arizona
Glen Jaross	Raytheon
Edward Kennelly	AER, Inc.
James Leitch	Ball Aerospace & Technologies Corp.
Jean-Luc Moncet	AER, Inc.
John Pickle	AER, Inc.
John Qu	Raytheon
Miroslav Predny	AER, Inc.
Hélène Rieu	AER, Inc.
Juan Rodriguez	Ball Aerospace & Technologies Corp.
James Russell	Hampton University
Courtney Scott	AER, Inc.
Thomas Swissler	Consultant

2 Scientific Basis of Infrared Ozone Retrievals

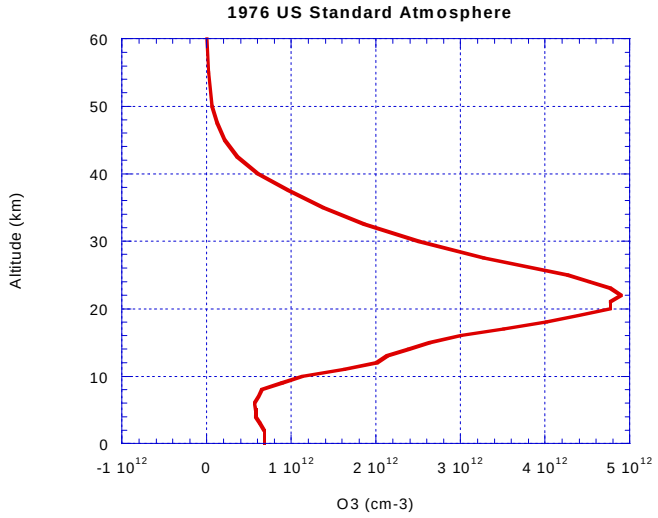
The retrieval of atmospheric ozone from space can be accomplished using sensors operating in the microwave, thermal infrared (IR), or UV/Visible regions of the radiometric spectrum. OMPS consists primarily of sensors operating in the UV/Visible range of the spectrum and designed to exploit heritage algorithms. However, this results in reduced coverage due to the inability to obtain useful data for large solar zenith angles and under nighttime conditions. To obtain estimates of the total column ozone for solar zenith angles greater than 80 degrees, we have chosen to use the IR radiances measured by the CrIS sensor. The spectral region chosen, roughly 1000 – 1065 cm^{-1}, is dominated by transitions of the ν_3 ozone fundamental, the strongest band of ozone in the infrared.

The retrieval of total column ozone amounts from IR radiances in an operational environment has a long heritage through the TIROS Operational Vertical Sounder (TOVS) program [Planet *et al.*, 1984]. In Section 2.1, following the work of Clough *et al.* [1995], we describe the physical basis of the ozone retrieval problem and provide rationale for why the CrIS ozone retrievals should provide higher total column ozone accuracy than the TOVS retrievals. Section 2.2 briefly describes the characteristics of the CrIS sensor, including comparisons to higher and lower spectral resolution designs (e.g. TES and TOVS). Section 2.3 outlines the retrieval algorithm (with the details of the forward and inverse models given in Section 3). Finally, Section 2.5 lists the major error sources impacting the ozone retrieval (with details given in Section 7).

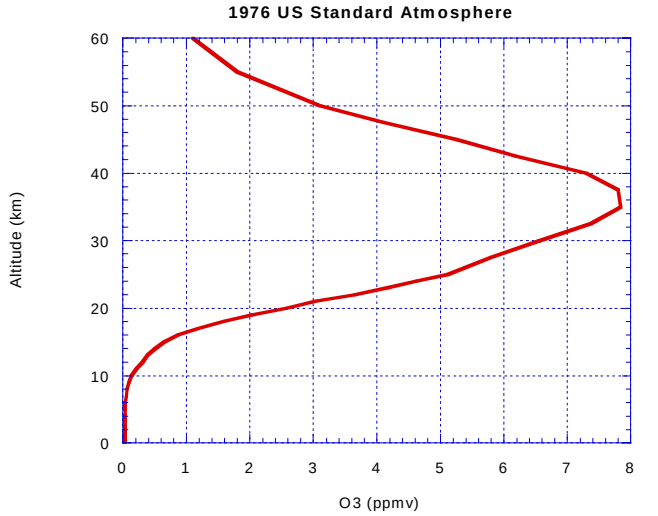
2.1 Physical Description

The ozone volume mixing ratio (vmr) and number density profiles corresponding to the U.S. Standard Atmosphere [Anderson *et al.*, 1986] are shown in Figure 2.1-1. Although the vmr decreases by 2.5 orders of magnitude between its maximum at 35 km and the surface, the variation of the number density, a quantity more directly related to the atmospheric radiative transfer, exhibits a significantly smaller decrease from its maximum at 22 km to the surface. A calculation of the brightness temperature response function (Figure 2.1-2) indicates that the largest sensitivity to ozone occurs around 200 mb (about 12 km for the U.S. Standard Atmosphere).

The retrieval of total column ozone from space has been demonstrated operationally by the TOVS program [e.g., Miller, 1989] using the High Resolution Infrared Radiation Sounder (HIRS). More recently there has been discussion about the possibility of obtaining tropospheric ozone amounts from space. The fact that the ozone number density in the troposphere is roughly 10% of that in the stratosphere is an important consideration in assessing the feasibility of retrieving tropospheric ozone. A second important consideration relates to the width and shape of the ozone lines as a function of altitude. The Doppler component of the line shape is dominant in the stratosphere. At these altitudes the half width of the ozone lines is of the order of 0.0015 cm^{-1} at 1000 cm^{-1} (see Clough *et al.*, 1992 for a complete description). Consequently the radiative effects of ozone at the higher altitudes are concentrated close to the line center. In the altitude regime below 22 km, collisional broadening (i.e. a Lorentzian line shape) is dominant with the half width increasing linearly with pressure to 0.08 cm^{-1} at the surface.



(a)



(b)

Figure 2.1-1. Vertical profile of ozone (a) number density and (b) volume mixing ratio from the surface to 60 km for the U.S. Standard Atmosphere (data from Anderson et al., 1986).

The OMPS algorithm uses ozone lines from the ν_3 ozone band, the strongest infrared ozone band. These spectral lines cover the full range of opacity, from linear absorption of the weak lines to full saturation for strong lines. For each of the individual lines the ozone absorption is strongest at line center and the emission is saturated at an altitude just above the stratopause (~50 km). The brightness temperature increases slightly away from line center corresponding to emission from lower in the atmosphere, near the stratopause, decreases sharply to a brightness temperature representative of the tropopause temperature, and then increases as the wing of the line follows the tropospheric temperature lapse rate toward the surface. The strong spectral

dependence close to line center is due to the Doppler component of the Voigt line profile, with the weaker spectral dependence in the wings attributable to the Lorentz line shape.

The above discussion indicates that the ability to retrieve ozone profile information is directly related to the spectral resolution of the measurement. To first order, if the resolution is such that the entire ozone line (from wing to wing) falls within a single channel of the sensor, the measurement will be sensitive only to changes in the entire ozone column. Conversely, if the spectral resolution is sufficient to resolve the line structure shown in Figure 2.1-3, much more profile information will be obtained. Because CrIS has a higher spectral resolution than HIRS one can expect that the ozone profile retrieval performance will be better than that of the TOVS ozone retrievals. However, the resolution is still a limiting factor in the ability to measure detailed profile information.

When considering the use of a particular spectral region for retrieving atmospheric parameters it is important to consider the overlap of these parameters with other atmospheric variables. For the 9.6 μm ozone band, only weak carbon dioxide hot band lines and a few high excitation lines of water vapor add significantly to the line opacity. Broad absorption is produced by the self-broadened water vapor continuum, bands of the more abundant halocarbons, particularly CFC-11 [McDaniel *et al.*, 1992], and aerosols. The impact of these parameters on the ozone retrieval can be mitigated if they are properly accounted for in the retrieval algorithm. The inclusion of these factors is discussed in more detail in Section 3.

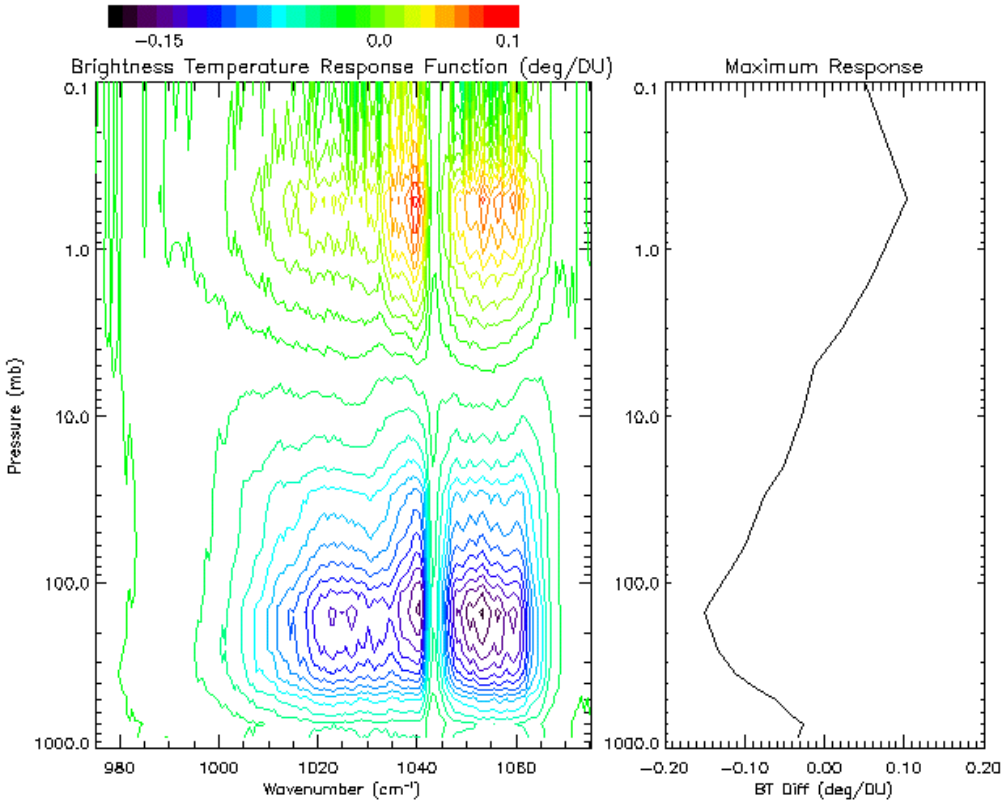


Figure 2.1-2. Brightness temperature response function for standard ozone profile.

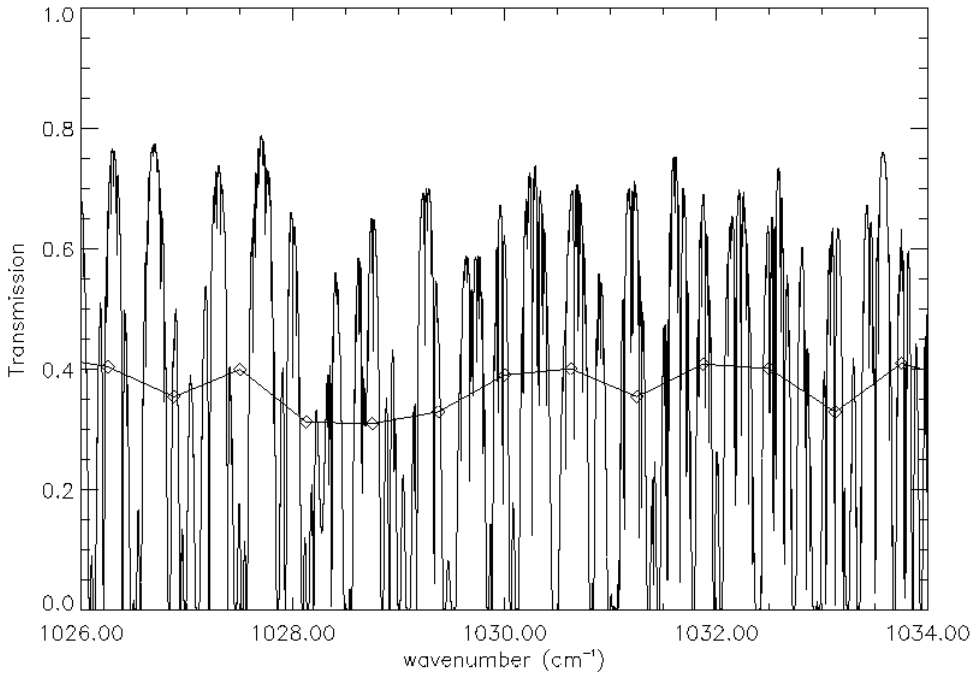


Figure 2.1-3. Spectral monochromatic transmission in the ozone band. The line connected by diamonds represents the spectrum convolved with a Hamming function at the CrIS spectral resolution (0.625 cm^{-1}), indicating the inability of the CrIS to separate the line structure.

The retrieval of ozone using infrared radiances has two key advantages over the use of UV/visible radiances. The obvious advantage is the ability to retrieve ozone at night, and it is for this reason we have chosen to develop IR algorithms. The other advantage is that IR detectors have negligible sensitivity to the high number of proton impacts which occur as the spacecraft travels through the South Atlantic Anomaly (SAA). In this region the IR retrievals will provide a co-located means of testing the validity of the UV/vis retrievals.

2.2 Sensor Characteristics

In the pre-PDR phase of this project, the final CrIS design was unknown and all IR-TCA studies were conducted based on the notional design parameters given in the CrIS Sensor Requirements Document (dated 2 September 1998). Trade studies were conducted to determine the suitability of the SRD parameters, and the impact of minor changes to these parameters, on the retrieval of total column ozone. Extreme changes from this design, such as a HIRS or TES sensor, were also investigated. Now that the ITT team has been selected for development of the CrIS sensor, all simulations are done using the current sensor specifications. An ICSR was submitted to the IPO (21 December 1999) in order to obtain the relevant sensor and algorithm documentation for the CrIS, and this data is now being used for all of the simulations. Of particular importance to the OMPS-IR algorithm were the field-of-view parameters, specification of the CrIS channels and bandwidth, and an up-to-date estimate for the sensor radiometric noise.

While specific requirements placed upon the CrIS sensor derived from our IR-TCA simulations are outlined in Section 5 of this document, the overall IR-TCA development assumes the characteristics outlined below. Note that some items are labeled "TBD" pending receipt of additional CrIS documentation. (For these parameters we currently assume SRD specifications).

1. The CrIMSS payload consists of both infrared (CrIS) and microwave components.
2. The microwave component includes sensors similar to Advanced Microwave Sounding Unit (AMSU) and the Microwave Humidity Sounder (MHS), both of which are cross-track scanning, multispectral microwave radiometers.
3. The NeDN and calibration radiometric bias errors for the microwave SDR data are known.
4. The CrIS is a cross-track scanning Michelson interferometer with a nadir footprint of 0.963 degrees, or about 14 km for an 833 km orbit.
5. The CrIS is divided into three spectral bands with the channels given in Table 2.2-1.
6. The CrIS is designed such that the CrIMSS will meet the vertical moisture profile EDR requirements, vertical temperature profile EDR requirements, and vertical pressure profile EDR requirements as given by the CrIS SRD.
7. The CrIS SDR data consists of a calibrated, apodized spectrum with well-known instrument line shape (ILS) with channel centers remapped to a fixed wavenumber grid.
8. The CrIS SDR calibration includes corrections for off-optical-axis effects such as spectral shifts, self-apodization, phase distortion, etc., with associated quality control flags.
9. The CrIS NeDN, radiometric bias, spectral shift errors, band-to-band co-registration errors, and jitter errors are known.
10. The NeDN is no larger than that given in (TBD CrIS document, currently private communications from ITT).
11. The three CrIS bands have a spatial co-registration to within TBD of the FOV diameter and the scan angle of the measurement is known.
12. The latitude and longitude for each CrIS FOV is known to within < 1.45 km mapping error.

Table 2.2-1. Nominal Spectral Band Characteristics for CrIS. The spectral resolution is given as the value computed from the optical path difference (i.e., 1/2L).

CrIS Band	Wavenumber Range	Spectral Resolution	Number of Channels
1 – LWIR	650-1095 cm^{-1}	0.625 cm^{-1}	713
2 – MWIR	1210-1750 cm^{-1}	1.25 cm^{-1}	433
3 – SWIR	2155-2550 cm^{-1}	2.50 cm^{-1}	158
		Total	1304

2.3 Retrieval Algorithm

The retrieval algorithm consists of a forward radiance model and a radiance inversion model. The forward model, described in Section 2.3.1, is used to compute radiances given various atmospheric parameters (e.g., temperature, water vapor and ozone profiles, surface properties, cloud properties, etc.) and compute the derivatives of radiance with respect to the particular parameters of interest. The inversion model, outlined in Section 2.3.2, is used to convert the

difference between measured radiances and modeled radiances into changes in various atmospheric parameters.

2.3.1 Forward Model

The radiative transfer model used to relate atmospheric parameters to the measured radiances is known as the forward model. This model is used to calculate the radiance measured at the top of the atmosphere. For a given monochromatic spectral point the radiance leaving the top of the atmosphere (TOA) at viewing angle θ may be written as:

$$R_{\text{thermal}}(\nu,\theta) = \int_{\ln p_s}^{\ln p_{TOA}} B[\nu,T(p)]\frac{\partial\tau(\nu,p,\theta)}{\partial \ln p}d\ln p + \varepsilon_s B[\nu,T_s]\tau(\nu,p_s,\theta) + (1-\varepsilon_s)\tau(\nu,p_s,\theta)\int_{\ln p_s}^{\ln p_{TOA}} B[\nu,T(p)]\frac{\partial\tau(\nu,p,\theta)}{\partial \ln p}d\ln p \quad (1)$$

where the Planck function emission for layer p at temperature T is given by B[ν,T(p)], τ(ν,p,θ) is the layer-to-space transmission, the subscript s refers to the surface, and ε_s and ρ_s are the surface emissivity and reflectivity. It should be noted that Equation (1) neglects atmospheric scattering and clouds, and assumes that the surface is Lambertian. In daytime situations a solar term must be added to Equation (1):

$$R_{TOA}(\nu,\theta) = R_{thermal}(\nu,\theta) + \rho_s S(\nu)\tau(\nu,p_s,\theta)\tau(\nu,p_s,\theta_{sun})\cos\varphi_{sun} \quad (2)$$

where the solar zenith and azimuth angles are given by θ_{sun} and φ_{sun}, and S(ν) is the solar brightness given by:

$$S(\nu) = 2.16x10^{-5}\pi B(\nu,T_{sun}) \quad (3)$$

The forward model used in the IR-TCA is based upon the optimal spectral sampling (OSS) method developed at AER. The OSS is a line-by-line (monochromatic) radiative transfer model that has been optimized for both speed and accuracy. The OSS can be applied to any portion of the spectrum, but the discussions herein are limited to those features relevant to the infrared. Comparisons of OSS with the industry-standard LBLRTM are given in Section 7.

Because the OSS method utilizes essentially a monochromatic approach to the radiative transfer, the gradient of the forward model with respect to all relevant atmospheric/surface parameters can be computed efficiently using an analytical scheme. Computation of radiances and derivatives with the OSS method uses a generic recursive scheme developed for the modeling of upward, downward-looking and limb-viewing instruments and used in atmospheric retrievals from CIRRIS-1A [Miller *et al.*, 1999]. Figure 2.3-1 defines the numbering conventions for the layered atmosphere. Level 0 represents the top of the atmosphere while level N represents the surface.

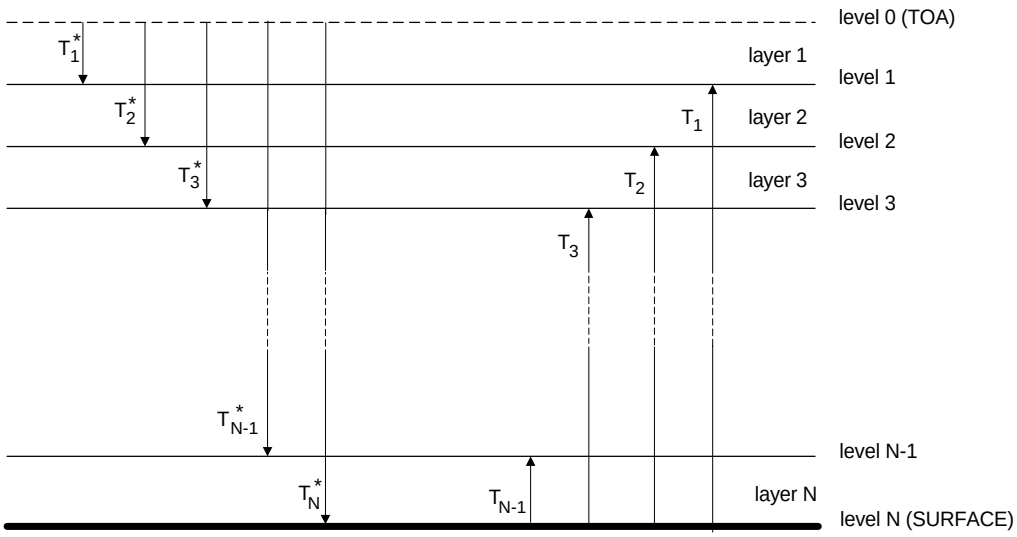


Figure 2.3-1. Schematic diagram showing the numbering convention for the atmospheric layers used by OSS. T refers to transmittance.

2.3.1.1 OSS Transmittance Module

To compute transmittance, the OSS model makes use of pre-stored monochromatic layer optical depths for the relevant atmospheric gases at the selected wavenumber locations. The gases are split into two groups, those that have a fixed molecular amount and those that are variable. Because a single optical depth represents the fixed gases this grouping reduces the storage requirements for the optical depth tables. For each species, the optical depths are stored at N temperatures for each pressure layer used in the discrete radiative transfer model. For each layer the temperature range $T(p_i)_{min}$ to $T(p_i)_{max}$ spans the temperatures expected for that layer based on the profiles in the TIGR [Chedin, *et al.*, 1985] and NOAA-88 atmospheric profile databases. Currently this temperature domain is sampled uniformly and the optical depths are stored at 10 temperatures. This scheme could be modified to better optimize this sampling based on the likelihood of various temperatures at a given level, rather than a uniform grid.

The optical depth tables are calculated using the LBLRTM radiative transfer model along with the molecular amounts from the US Standard Atmosphere profile [Anderson *et al.*, 1986]. This model contains state-of-the-art physics, including the latest version of the Clough-Kneizys-Davies (CKD) water vapor continuum [Clough *et al.*, 1989] and spectral line parameter information from the HITRAN database [Rothman *et al.*, 1998]. Because of the CKD formulation the method of optical depth calculation must be modified slightly for water vapor, since the self-broadened component of the water vapor continuum contains a quadratic dependence on the number density. As such the self-broadened component is separated from the water lines and the foreign-broadened component of the continuum.

Currently the fixed gases are defined as CO_2, N_2O, CH_4, CO and O_2, while H_2O and O_3 are the variable species. While this is sufficient for the current simulations, the next version of the code will have only O_2 as the fixed gas, to allow for the global variation of CO_2, N_2O and CH_4. For a given layer each optical depth is linearly interpolated to the layer temperature. If the layer

pressure is given by p_l and the temperature by θ_l, the optical depth interpolation for the jth species may be written as

$$\tau_j(p_l,\theta_l) = \tau_j(p_l,\theta_a) + \frac{\tau_j(p_l,\theta_b) - \tau_j(p_l,\theta_a)}{\theta_b - \theta_a}(\theta_l - \theta_a) \tag{4}$$

The correct optical depths for the variable species are then obtained by multiplying the temperature-interpolated optical depth by the ratio of the layer amount to the standard amount. This is equivalent to multiplying the species' molecular absorption coefficient by the molecular amount, but reduces numerical accuracy problems encountered because of the many orders of magnitude difference in the value of the absorption coefficient and the molecular amount. Of minimal impact, and thus neglected in the current formulation, is the difference in the self-broadened component of the line shape between the standard density used to compute the stored optical depths and the actual layer density. Mathematically the calculation of the total optical depths τ_l at p_l and θ_l for a single monochromatic point can be written as follows:

$$\begin{aligned}\tau_l = {} & \tau_{fix}(p_l,\theta_l,\alpha_{fix,std}) + \tau_{O3}(p_l,\theta_l,\alpha_{O3,std})\left[\frac{\alpha_{O3,l}}{\alpha_{O3,std}}\right] \\ & + \tau_{H2O}(p_l,\theta_l,\alpha_{H2O,std})\left[\frac{\alpha_{H2O,l}}{\alpha_{H2O,std}}\right] + \tau_{H2O,self}(p_l,\theta_l,\alpha_{H2O,std})\left[\frac{\alpha_{H2O,l}}{\alpha_{H2O,std}}\right]^2\end{aligned} \tag{5}$$

where α represents the molecular path amount (molecules/cm^2), *fix* the fixed gas components, and *std* standard atmosphere concentrations and conditions.

Table 2.3-1 summarizes typical storage requirements for the optical depth tables. The current set of optical depths is stored at each monochromatic spectral point for 39 layers and 10 temperatures. While the three CrIS bands have spectral resolutions of 0.625, 1.25 and 2.5 cm^{-1}, these optical depth tables were generated for spectral resolutions of 0.04, 0.08, and 0.16 cm^{-1}. This was done for convenience during the spectral resolution trade studies, as the monochromatic spectral locations for the lower spectral resolution cases will always be a subset of the higher spectral resolution data set. However, prior to running the radiative transfer model, the size of these files is reduced by (a) eliminating duplicate spectral points, (b) restricting the bandwidth of the file to the bandwidth of the channels to be used in the radiative transfer calculation, and (c) restricting the coefficients to those used by the sensor channels.

Table 2.3-1. Storage Requirements for Monochromatic Optical Depth Tables Prior to Compression

Band	Spectral Resolution	Spectral Band	Number of Points	Total Size
1	0.04 cm^{-1}	550 – 1149.96 cm^{-1}	47331	~73.84 Mbytes
2	0.08 cm^{-1}	1150 – 1799.92 cm^{-1}	20072	~31.31 Mbytes
3	0.16 cm^{-1}	1800 – 2999.84 cm^{-1}	24455	~38.15 Mbytes
		Total	91858	~143.3 Mbytes

The size of the compressed optical depth tables is given in Table 2.3-2 for the nominal CrIS bandpass and spectral resolution. The total of 15.56 Mbytes is for each of the optical depth tables (one file each for fixed gases, ozone, water lines, and water self-continuum), resulting in a total storage requirement of approximately 62.25 Mbytes. The storage requirement can be further reduced by noting that in the infrared, pressure broadening dominates in the atmosphere below 10 mb. Thus a single optical depth table entry will suffice for all layers above 10 mb. With the current selection of layering, this results in a reduction of 9 layers, for total savings of about 14.4 Mbytes. As mentioned above, further optimization of the number of temperatures for which the optical depths are stored at each layer would also reduce the storage requirements.

Table 2.3-2. Storage Requirements for Optical Depth Tables After Bandwidth Reduction and Elimination of Duplicate Spectral Points

Band	Spectral Resolution	Spectral Band	Number of Points	Total Size
1	0.625 cm^{-1}	635 – 1095 cm^{-1}	4812	~7.51 Mbytes
2	1.25 cm^{-1}	1210 – 1750 cm^{-1}	3664	~5.72 Mbytes
3	2.50 cm^{-1}	2155 – 2800 cm^{-1}	1491	~2.33 Mbytes
		Total	9967	~15.56 Mbytes

2.3.1.2 OSS Radiance Module

2.3.1.2.1 Overview

In clear conditions, OSS computes monochromatic (R_ν) radiances using a discrete form of the radiative transfer equation (RTE):

$$R_\nu = \sum_{i=1}^{N-1} \left(T_{\nu,i} - T_{\nu,i+1}\right) B_\nu^+(\theta_i) + \varepsilon_{\nu,s} T_{\nu,N} B_\nu^+(\theta_s) + \left(1-\varepsilon_{\nu,s}\right) T_{\nu,N} \sum \left(T_{\nu,i+1}^* - T_{\nu,i}^*\right) B_\nu^-(\theta_i) \quad (6)$$

where θ_l is the temperature of layer *l*, $B_\nu^{\pm}$ represents the upward and downward Planck emission of the layer at wavenumber ν, and $\varepsilon_{\nu s}$ is the surface emissivity. The convention for layer numbers and transmittance (T, T*) is given in Figure 2.3-1. Derivatives of R_ν with respect to constituent concentration or temperature in layer *l* are obtained by differentiating (6),

$$\frac{\partial R}{\partial X_l} = -\frac{\partial \tau_l}{\partial X_l}\left\{\begin{array}{l}\left[-T_{l+1}B_l + \sum_{i=l+1}^{N-1}(T_i - T_{i+1})B_i + T_N\varepsilon_s B_s + (1-\varepsilon_s)T_N\sum_{i=1}^{N-1}(T_{i+1}^* - T_i^*)B_i\right]\sec\theta_{obs} \\ -\left[(1-\varepsilon_s)T_N T_l^* B_l + (1-\varepsilon_s)T_N\sum_{i=1}^{l+1}(T_{i+1}^* - T_i^*)B_i\right]\sec\theta_{down}\end{array}\right\} \quad (7)$$
$$+\frac{\partial B_l^+}{\partial X_l}(T_l - T_{l+1}) + \frac{\partial B_l^-}{\partial X_l}(T_{l+1}^* - T_l^*)$$

where θ_{obs} and θ_{down} are the observation and downwelling path angles, respectively, X_l stands for either the layer temperature, θ_l or the concentration of the m[th] constituent, ω_l^m.

With the exception of the handling of the surface terms, the OSS utilizes a recursive procedure for the integration of the RTE and calculation of the derivatives. This method exploits the fact that a perturbation in temperature or constituent concentration within any given layer of the atmosphere does not affect the emission in the atmospheric slab between this layer and the observer. Therefore, derivatives can be obtained at low cost if the RTE is integrated by adding layers sequentially in the direction of the observer. The procedure is more apparent by introducing the quantities Σ_l^- and Σ_l^+ defined as the contribution to the observed radiance of the downward emission (reflected at the surface) from the atmosphere above level *l* and the contribution of the atmosphere below level *l* plus reflected downward radiation, i.e.,

$$\Sigma_l^- = \sum_{i=1}^{l}(T_i' - T_{i-1}')B_i^- \quad (8)$$

and

$$\Sigma_l^+ = \sum_{i=l+1}^{N}(T_{i-1} - T_i)B_i^+ + \varepsilon_s T_N B_s + \sum_{i=1}^{N}(T_i' - T_{i-1}')B_i^- \quad (9)$$

where the two-path attenuation from level *l* to the surface and then upward to space is given by:

$$T_l' = (1-\varepsilon_s)T_N T_l^* \quad (10)$$

Using the definitions of (8-10), (6) and (7) become,

$$\frac{\partial R}{\partial X_l} = \left[\frac{\partial R}{\partial X_l}\right]_u + \left[\frac{\partial R}{\partial X_l}\right]_d$$
$$= \left[\frac{\partial \tau_l}{\partial X_l}\left(-\Sigma_{l-1}^{-} + B_l^{-}\mathrm{T}'_{l+1}\right) + \frac{\partial B_l^{-}}{\partial X_l}\left(\mathrm{T}'_{l+1} - \mathrm{T}'_l\right)\right] + \left[\frac{\partial \tau_l}{\partial X_l}\left(-\Sigma_{l+1}^{+} + B_l^{+}\mathrm{T}_l\right) + \frac{\partial B_l^{+}}{\partial X_l}\left(\mathrm{T}_l - \mathrm{T}_{l-1}\right)\right] \quad (11, 12)$$

In the current version of the IR-OSS algorithm, the following simplifying assumptions are made in the calculation of the derivatives:

1) Atmospheric transmittances do not depend on temperature.
2) Layer emission is computed as the average Planck emission for the layer, $B_l^{+} = B_l^{-} = \overline{B}_l = \left(B_l^{l} + B_l^{u}\right)/2$, and does not depend on layer optical depth.

The second approximation is adequate as long as layers are not optically thick, i.e., the vertical pressure grid is sufficiently fine and no cloud is present.

2.3.1.2.2 Practical Implementation

In the first pass through the atmosphere, at any given wavenumber, the algorithm computes the profile of transmittance from space. A recursive procedure for the computation of radiances and analytical derivatives then follows directly from (12):

1) Initialization: set $\Sigma_0^{-} = 0$

2) If $\mathrm{T}_N > 10^{-4}$, add layers successively from TOA down to surface. Update Σ_l^{-} at each step and compute the first part of radiance derivatives:

$$\Sigma_l^{-} = \Sigma_{l-1}^{-} + \left(\mathrm{T}'_l - \mathrm{T}'_{l-1}\right)\overline{B}_l \quad (13)$$

$$\left[\frac{\partial R}{\partial \tau_l}\right]_d = -\Sigma_{l-1}^{-} + \overline{B}_l\ \mathrm{T}'_{l-1} \quad (14)$$

$$\left[\frac{\partial R}{\partial \Theta_l}\right]_d = \frac{\partial \overline{B}_l}{\partial \Theta_l}\left(\mathrm{T}'_l - \mathrm{T}'_{l-1}\right) \quad (15)$$

3) Add the surface term and compute the radiance derivatives with respect to surface emissivity and temperature:

$$\Sigma_N^{+} = \Sigma_N^{-} + \mathrm{T}_N \varepsilon_s B_s \quad (16)$$

$$\frac{\partial \mathrm{R}}{\partial \Theta_s} = \mathrm{T}_N \varepsilon_s \frac{\partial \mathrm{B}_s}{\partial \Theta_s} \quad (17)$$

4) Update Σ_l^+ by adding layers from the surface up to the TOA and compute the second part of derivatives:

$$\frac{\partial R}{\partial \tau_l} = -\Sigma_l^+ + \overline{B}_l \, \mathrm{T}_l + \left[\frac{\partial R}{\partial \tau_l}\right]_d \tag{18}$$

$$\frac{\partial R}{\partial \Theta_l} = \frac{\partial \overline{B}_l}{\partial \Theta_l}\left(\mathrm{T}_{l-1} - \mathrm{T}_l\right) + \left[\frac{\partial R}{\partial \Theta_l}\right]_d \tag{19}$$

$$\Sigma_{l-1}^+ = \Sigma_l^+ + \left(\mathrm{T}_{l-1} - \mathrm{T}_l\right)\overline{B}_l \tag{20}$$

5) Set $R_\nu = \Sigma_0^+$ and compute the derivatives with respect to temperature and layer amounts for all molecular species:

$$\frac{\partial R_\nu}{\partial \omega_l^m} = \frac{\partial R_\nu}{\partial \tau_{\nu,l}} \times k_{\nu,l}^m; \quad m = 1,\ldots,M \tag{21}$$

2.3.1.2.3 Solar Radiation

In addition to the atmospheric (thermal) component of radiation reaching the sensor, the algorithm must also account for solar radiation that is transmitted through the atmosphere and reflects off of clouds or the surface. The solar radiance reaching the sensor is given by:

$$\mathrm{R}_{\mathrm{solar}} = \rho_s \mathrm{T}_N \exp\left(-\sum_l \tau_l^o \sec\theta_{\mathrm{solar}}\right) \mathrm{F}_o \tag{22}$$

$$\frac{\partial \mathrm{R}_{\mathrm{solar}}}{\partial \tau_l^o} = -\left(\sec\theta_{\mathrm{obs}} + \sec\theta_{\mathrm{solar}}\right)\mathrm{R}_{\mathrm{solar}} \tag{23}$$

The OMPS-IR forward model currently uses the monochromatic, full disk, extraterrestrial solar irradiance spectrum from Kurucz [1992]. This spectrum is based on solar photospheric modeling and has been validated in part by interferometer data from Kitt Peak Observatory. Details regarding the inclusion of the solar term in the OMPS-IR algorithm are given in AER AIPT memo AER-AIPT-006.

2.3.2 Inversion Model

A variety of approaches ranging from physical to statistical methods exist for the retrieval of geophysical parameters from remotely sensed observations [Houghton *et al.*, 1984]. The approach we have taken in developing an IR retrieval algorithm for OMPS is a version of the method of nonlinear least squares (alternatively referred to as the method of minimum variance). This method is particularly well suited to the OMPS retrieval problem because the number of observed spectral radiance values (over 1000) is large in relation to the number of parameters in

the profile to be retrieved (less than 100). The following discussion follows that of Clough *et al.* [1995] and Rodgers [1976, 1990].

The unconstrained nonlinear least squares problem may be expressed in terms of an objective function $\phi(x)$ as the l_2 norm of the spectral residuals,

$$\phi(\mathbf{x}) = \sigma^2(\mathbf{x}) = \frac{1}{2}\left\|\mathbf{F}(\mathbf{x}) - \mathbf{R}\right\|_2^2 \tag{24}$$

where $\sigma^2(x)$ is the variance, $\mathbf{F}(x)$ is the forward model function dependent on the atmospheric state parameter $\mathbf{x}$, and $\mathbf{R}$ is the measured spectral radiance function. Issues of weighting and normalization will be discussed later. In the method of least squares, a solution is sought for the state function $\mathbf{x}$ at which the objective function is a minimum. Given an initial $\mathbf{x}$, we seek the change in state $\delta\boldsymbol{x}$ for which the first derivative of the objective function is zero.

Expanding $\boldsymbol{\phi}'(\mathbf{x} + \delta x)$ in a Taylor series, we have for the minimum of $\boldsymbol{\phi}(\mathbf{x})$ the result that

$$\phi''(\mathbf{x})\delta\mathbf{x} = -\phi'(\mathbf{x}) \tag{25}$$

where the derivatives of the objective function are given by

$$\phi'(\mathbf{x}) = \left[\nabla\mathbf{F}(\mathbf{x})\right]\left[\mathbf{F}(\mathbf{x}) - \mathbf{R}\right] \tag{26}$$

$$\phi''(\mathbf{x}) = \nabla\mathbf{F}(\mathbf{x})^T\nabla\mathbf{F}(\mathbf{x}) + \left[\nabla^2\mathbf{F}(\mathbf{x})\right]\left[\mathbf{F}(\mathbf{x}) - \mathbf{R}\right] \tag{27}$$

The result may be expressed as

$$\left\{\nabla\mathbf{F}(\mathbf{x})^T\nabla\mathbf{F}(\mathbf{x}) - \left[\nabla^2\mathbf{F}(\mathbf{x})\left\{\mathbf{R} - \mathbf{F}(\mathbf{x})\right\}\right]\right\}\delta\mathbf{x} = \nabla\mathbf{F}(\mathbf{x})\left\{\mathbf{R} - \mathbf{F}(\mathbf{x})\right\} \tag{28}$$

Care has been taken here to retain terms of the same order to properly address the nonlinear problem. For the linear least squares problem, the second-derivative term containing $\boldsymbol{\nabla^2\mathbf{F}(\mathbf{x})}$ is zero. If all terms are retained from (28) in the solution for $\delta\mathbf{x}$, the method is referred to as the Newton method; if the second-derivative term is neglected, the method is known as the Gauss-Newton method. Extensive consideration of these and related methods for the iterative solution for the state vector $\boldsymbol{x}$ appears in the optimization literature [e.g., Gill et al., 1981].

For many applications all elements of the residual function, $[\mathbf{R} - \mathbf{F}(\mathbf{x})]$, are small at the solution point with the consequence that the second-derivative term may be neglected. In the present application, for which a large number of spectral elements will be used, each with a potentially significant value for the residual even at the solution point, the contribution from the second-derivative may not be ignored. The diagonal elements of $\boldsymbol{\nabla^2\mathbf{F}(\mathbf{x})}$ with respect to parameter space are amenable to calculation through the use of symmetric finite differences, but the off-diagonal elements appear to be prohibitively expensive to calculate given the computational cost of the line-by-line radiative transfer model. For some problems it proves adequate to use only the diagonal elements. However, for the present case, the resulting matrices are not even assured to

be positive definite. A preliminary assessment of the role of the second-derivative term for the ozone retrieval problem confirms that the principal effect of its neglect is to slow the rate of convergence to the solution and to increase the possibility of converging to a secondary local minimum. Both of these issues are addressed with a different approach as discussed in the section on the retrieval from simulated radiance data enabling the effective application of the Gauss-Newton method.

At this stage in our development of the inverse problem it is useful to introduce the concept of a cost function. This function has two principal roles: (1) to control the direction and size of the step in the iterative solution of the nonlinear problem and (2) to introduce *a priori* knowledge into the solution. In contrast to many inverse problems associated with the retrieval of atmospheric state parameters, the current problem has the appearance of being over-determined. However, there is considerable redundancy in the observations (multiple channels), so that although the present case is not ill-posed, it might well be described as poorly posed. To address the solution of the nonlinear problem and to develop a formalism to incorporate *a priori* knowledge, it proves useful to augment the objective function with a penalty function $g(\mathbf{x})$, so that (24) becomes

$$\phi(\mathbf{x}) = \sigma^2(\mathbf{x}) + \mathbf{g}(\mathbf{x}) = \frac{1}{2}\left\|\mathbf{F}(\mathbf{x}) - \mathbf{R}\right\|_2^2 + \mathbf{g}(\mathbf{x}) \tag{29}$$

with the result that the equation for the minimized value of the augmented function is

$$\left\{\nabla\mathbf{F}(\mathbf{x})^{\mathrm{T}}\nabla\mathbf{F}(\mathbf{x}) - \left[\nabla^2\mathbf{F}(\mathbf{x})\left\{\mathbf{R} - \mathbf{F}(\mathbf{x})\right\}\right] + \mathbf{g}''(\mathbf{x})\right\}\delta\mathbf{x} = \nabla\mathbf{F}(\mathbf{x})\left\{\mathbf{R} - \mathbf{F}(\mathbf{x})\right\} + \mathbf{g}'(\mathbf{x}) \tag{30}$$

From Equation (30) it is convenient to introduce a more compact matrix notation, retaining essential consistency with that of Rodgers [1976], and express the change in $\mathbf{x}$ as an iterative update on x^n,

$$\mathbf{x}^{n+1} = \mathbf{x}^n + \left\{\mathbf{K}^{\mathrm{T}}\mathbf{W}\mathbf{K} - \mathbf{L}^{\mathrm{T}}\mathbf{W}\left[\mathbf{R} - \mathbf{F}(\mathbf{x})\right] + \mathbf{\Gamma}'\right\}^{-1}\left\{\mathbf{K}^{\mathrm{T}}\mathbf{W}\left[\mathbf{R} - \mathbf{F}(\mathbf{x})\right] + \mathbf{\Gamma}'\right\} \tag{31}$$

where $\mathbf{K}_{i,j}$ is the element of the Jacobian of the forward model associated with the i^{th} spectral radiance element and the j^{th} element of the state vector $\mathbf{x}$,

$$\mathbf{K}_{i,j} = \frac{\partial\mathbf{F}_i}{\partial\mathbf{x}_j} \tag{32}$$

$\mathbf{L}_{i,jk}$ is the element of the matrix associated with the second derivative of $\mathbf{F(x)}$ with respect to the j and k elements of $\mathbf{x}$,

$$\mathbf{L}_{i,j,k} = \frac{\partial^2\mathbf{F}_i}{\partial\mathbf{x}_j\partial\mathbf{x}_k} \tag{33}$$

Γ″and **Γ'** are a matrix and a vector, respectively, associated with the derivatives of the penalty function $g(\mathbf{x})$, and **W** is the weight matrix associated with the measurements. In the present work we take the generalized weight matrix to be the inverse of the error covariance matrix of the spectral radiance measurements, S_m,

$$\mathbf{W} = \mathbf{S}_m^{-1} \tag{34}$$

In the following development we suppress the second-derivative term. For the OMPS IR-TCA, the Jacobian is computed directly from analytical expressions, rather than using the more computationally expensive method of finite differences.

For the aspect of the problem in which the penalty function is used to control the direction and size of step, we follow the customary approach by defining $g(\mathbf{x})$ to be a quadratic function around the current state vector $\mathbf{x}^n$, such that the departure of **x** from $\mathbf{x}^n$ causes a quadratic increase in $\phi(\mathbf{x})$,

$$g(\mathbf{x}) = \frac{1}{2}\sum_i \boldsymbol{\gamma}_i (\mathbf{x} - \mathbf{x}^n)_i^2 \tag{35}$$

which has derivatives given by

$$\mathbf{g}'(\mathbf{x}) = \gamma(\mathbf{x} - \mathbf{x}^n) \tag{36}$$

$$\mathbf{g}''(\mathbf{x}) = \gamma \tag{37}$$

Equation (31), with the derivatives evaluated at $\mathbf{x}^n$, results in the well-known Levenberg-Marquardt method [Levenberg, 1944; Marquardt, 1963]. If **x** is updated iteratively, the expression can be written as

$$\mathbf{x}^{n+1} = \mathbf{x}^n + \left\{\mathbf{K}^T\mathbf{S}_m^{-1}\mathbf{K} + \boldsymbol{\gamma}\mathbf{I}\right\}^{-1}\left\{\mathbf{K}^T\mathbf{S}_m^{-1}\left[\mathbf{R} - \mathbf{F}(\mathbf{x})\right]\right\} \tag{38}$$

In Equation (38) the size and direction of step constrained through the values of γ_i. At the solution point, the difference between $\mathbf{x}^{n+1}$ and $\mathbf{x}^n$ is small such that γ_i has negligible impact. In many algorithms, the value of γ_i is relaxed toward zero as **x** approaches the solution point.

The error covariance of the retrieved state vector, S_x, is given by

$$\mathbf{S}_x = \left(\mathbf{K}^T\mathbf{S}_m^{-1}\mathbf{K}\right)^{-1} \tag{39}$$

in which we have assumed that the only error is due to measurement error, S_m. In this context a broad range of methods have been developed for nonlinear optimization as discussed by Gill *et al.* [1981].

At this point *a priori* information has not been used in the solution of the inverse problem. However, the present formalism allows this to be easily incorporated by assuming the existence of a proper *a priori* state vector $\mathbf{x}^a$ with associated error covariance matrix $\mathbf{S}_a$ [Marks and Rodgers, 1993; Rodgers, 1976, 1990; Backus and Gilbert, 1970]. In this case the equation for the update of **x** is obtained as

$$\mathbf{x}^{n+1} = \mathbf{x}^n + \left\{\mathbf{K}^T\mathbf{S}_m^{-1}\mathbf{K} + \mathbf{S}_a^{-1}\right\}^{-1}\left\{\mathbf{K}^T\mathbf{S}_m^{-1}\left[\mathbf{R} - \mathbf{F}(\mathbf{x})\right] + \mathbf{S}_a^{-1}(\mathbf{x}^a - \mathbf{x}^n\right\} \tag{40}$$

This approach is known as the maximum likelihood technique, or optimal estimation approach. In the simple case, $\mathbf{S}_a$ is assumed diagonal with inverse elements given by γ. The important point here is that at each iteration **x** is constrained toward the *a priori* state with its strength dependent on the error covariance matrix associated with that state. For the maximum likelihood method, the error covariance matrix associated with the retrieved state **x** is given by [e.g., Rodgers, 1990]

$$\mathbf{S}_x = \left(\mathbf{K}^T\mathbf{S}_m^{-1}\mathbf{K} + \mathbf{S}_a^{-1}\right)^{-1} \tag{41}$$

It should be noted that if the second-derivative term is available the error covariance of **x** takes the form:

$$\mathbf{S}_x = \left\{\mathbf{K}^T\mathbf{S}_m^{-1}\mathbf{K} - \mathbf{L}^T\mathbf{W}\left[\mathbf{R} - \mathbf{F}(\mathbf{x})\right]\mathbf{L} + \mathbf{S}_a^{-1}\right\}^{-1} \tag{42}$$

$$\mathbf{f}_j = \left\{\left[\mathbf{K}^T\mathbf{S}_m^{-1}\mathbf{K} + \mathbf{S}_a^{-1}\right]^{-1}\left[\mathbf{K}^T\mathbf{S}_m^{-1}\mathbf{K}\right]\right\}_{jj} \tag{43}$$

thus explicitly including a contribution dependent on the nonlinear aspect of the problem. It is useful to define an index indicating the fraction of the variance of the j^{th} element of **x** that is explained by the measurement that in the linear approximation is given by (43). This quantity is the diagonal element of the averaging kernel discussed by Rodgers [1990].

For a problem with a given experimental design that is ill-posed, the utilization of *a priori* information is a necessity. However, the assumptions critical to the maximum likelihood method should be well understood and fully considered for each situation. These are (1) that there exists a state $\mathbf{x}_a$, that is the valid mean of the ensemble of observed states of interest and (2) that there exists a normal distribution of states about $\mathbf{x}_a$ that can be represented by the error covariance matrix $\mathbf{S}_a$. With respect to the first point, it is clear that for studies of long-term trends, the bias introduced by using *a priori* information based on past observations can be a serious liability. With respect to the second point, it is not at all obvious that the ozone profiles, particularly in the troposphere, can be expected to follow a normal distribution on the spatial and temporal scales of interest. One can perhaps envision background ozone profiles that fluctuate quasi-randomly about a mean; however, for air masses affected by anthropogenic influences the existence of a useful mean with an associated error covariance matrix is a highly unlikely prospect. The objective of the retrieval method in the present application is to obtain the most likely profile for the atmospheric state associated with each measurement; attaining an ensemble mean consistent with *a priori* information does not meet this objective.

Having expressed a concern with respect to the maximum likelihood method, it is still important to pursue the introduction of *a priori* information in a less restrictive manner. It is well known that in solutions to the inverse problem with noise there generally arises an alternation in state parameter magnitude often referred to as jackknifing. The magnitude of this high-frequency component in the retrieval vector is clearly inconsistent with physical processes. To constrain this effect, it has been suggested that the first derivatives be constrained toward values more consistent with *a priori* information [e.g., Twomey, 1977]. It is straightforward to develop a penalty function $g(\mathbf{x})$ to attain this objective. In current least squares applications a more effective and less restrictive approach has been taken in which the second derivatives are constrained toward those consistent with the *a priori* profile. Although this method has not been used in the results presented in the present study, it is likely to play a role in the operational retrieval method to be developed to meet the OMPS objectives.

2.4 Clouds

2.4.1 Introduction

An important component of remote sensing algorithms is the accurate treatment of the effects of clouds, both in the measurement simulation and in the retrieval. Clouds are ubiquitous, and the extent to which they interfere with the desired measurement depends on the altitude and thickness of the cloud, the spatial extent relative to the sensor field-of-view, and the spectral measurement characteristics of the sensor. For example, microwave sensors can be less sensitive to the presence of clouds than visible wavelength sensors, while the physical characteristics of the cloud (e.g. cloud top altitude and particle size/type) will have different impacts for visible and infrared sensors. In this section we briefly describe the cloud characteristics that will be encountered by the OMPS-IR ozone algorithm, with emphasis on the Polar cases. We also describe the methods which will be employed to mitigate the impact of clouds on the ozone retrieval, and their implementation in the OMPS-IR algorithm. The test cases and expected performance of the OMPS-IR retrieval are described in Section 7.2.1.1.5.

Arctic clouds tend to be optically thin, low lying hazes and fogs (Curry et al., 1996). Because these have only a minimal radiometric signature, there is very little contrast between clear FOVs and nearby cloudy FOVs. The dominant feature of the temperature and moisture fields is the presence of near-surface inversions. These conditions are relatively stable and may persist for weeks, thus decoupling the surface from the lower troposphere. The challenge in validating any sort of algorithm used to mitigate the effect of clouds is in the difficulty of providing accurate ground-truth data. While part of this is due to the remote nature of the Polar regions, Polar night provides an additional challenge because lack of illumination and the presence of ice fogs. These conditions usually lead to an underestimation of the average cloud cover by surface observations, particularly for middle- and high-level clouds (Hahn et al., 1995). Alternatively, evidence of a systematic overestimation of winter cloud amounts by satellite sounders relative to surface observations suggests that the satellite sensors are sensitive to clear-sky ice particle precipitation, which is difficult to observe and is not technically classified as a cloud by surface observers (Wilson et al., 1993).

Calculation of performance estimates for the OMPS-IR algorithm under cloudy conditions requires knowledge of both the types of clouds present and their temporal and spatial variability. The morphology of Polar clouds is discussed in detail by Curry et al. (1996) and will not be repeated here. However, it is important to note that the annual variability of total cloud amount is dominated by low-level clouds, though it is not uncommon for summertime stratus clouds to occur in as many as five well-defined layers.

A further complication for the simulation of clouds, and ultimately for the mitigation of their effect on the retrieval algorithm, is the presence of highly variable surface characteristics (Kay and Barry, 1989). Surface conditions which present particularly difficult problems occur as the atmosphere is warming during the Polar summer. At this time the underlying surface can vary dramatically both spatially and temporally due to the movement of sea ice, the melting of snow and formation of fresh water or brackish ponds, and rapid variations in the temperature along the coastlines. These conditions result in sharp changes in the measured radiances, both within a specific FOV and between adjacent FOVs. Accurate simulation of algorithm performance requires that these conditions be considered.

Several studies have been conducted in which cloud statistics are derived from satellite measurements and, to some extent, validated through ground-based observations. Wylie reports on cirrus cloud statistics using HIRS data (Wylie et al., 1994 and 1999). In these studies it was found that there is a large seasonal change in cirrus cloud cover over Antarctica, with few clouds reported at any altitude in the austral winter. Schweiger et al. (1999) confirms a poor match between surface-observed and satellite-derived cloud amounts during the winter, but very good agreement in summer months. They also show that the mean monthly cloud amounts for the area north of 80° N are about 60% for October through April and rise to about 80% for June through September.

Cloud cover was also examined during the FIRE-ACE campaign of April through August 1998 as a part of SHEBA (Curry et al., 1999). In particular, Minnis et al. (1999) examined AVHRR data in conjunction with ground-based measurements. Their average cloud fractions agreed with the results listed above from Schweiger. Cloud top altitude was also examined in this study, and the results show that the altitudes of multiple layers of clouds are difficult to validate because ground-based observations may be limited to the lower cloud altitude while the satellite derives the altitudes of the upper cloud layers. In the final analysis, about 25% of the clouds are found below 500 m, and 50% are below 2 km. The remaining clouds are generally between 2 and 6 km. This is generally true over the entire observation period, except for July which tended to have more extremely low and very high (> 4 km) clouds observed than during the other months.

2.4.2 Overview of Cloud Mitigation Approaches

Mitigation of the effects of clouds on the retrieval of atmospheric parameters can be accomplished in a number of ways using threshold methods, various radiative transfer modeling techniques, and via statistical relationships (Key and Barry, 1989; Rossow et al., 1989). In the most simple approach, a cloud detection scheme can be employed to determine cloudy FOVs and eliminate them from the retrieval process (this is referred to as the “hole-hunting” technique). While this technique will in general simplify the retrieval algorithms, the large extent of cloud cover results in the loss of a significant amount of data, except for sensors with very high spatial

resolution (Smith et al., 1996). A technique similar to the hole-hunting approach is an algorithm which detects overcast FOVs and then performs a retrieval down to the cloud top. This approach is somewhat more complex in that it requires accurate knowledge of the cloud top pressure, and can lead to large errors if the cloud top is incorrectly identified. At the other extreme of complexity is an algorithm which simultaneously retrieves the relevant cloud parameters, e.g. the simultaneous cloud parameter retrieval ("SCPR", briefly described in the CrIS ATBD). This type of approach is applicable for sensors which have sufficient spatial and spectral information (and low noise characteristics) to separate the cloud information from that of the surface and other atmospheric constituents.

A compromise between these cloud mitigation techniques is the "cloud-clearing" approach, which has been selected as the baseline approach for the CrIS EDR algorithm (CrIS ATBD, 2000). Based on the work of Smith (1968) and Chahine (1974, 1977), cloud-clearing is a physical approach which blends the cloud-detection schemes of the hole-hunting technique with the cloud parameter retrievals of SCPR by exploiting the radiometric differences between adjacent FOVs to minimize the effect of clouds. With the cloud-clearing approach the algorithm is designed to compare multiple FOVs (or clusters of FOVs) to compute the radiance that would be observed had the FOV been clear, thus requiring *N+1* FOVs for the determination of *N* cloud formations. The biggest shortcoming of this approach is the assumption that the only source of spatial inhomogeneity is due to clouds. This can lead to errors in the retrieval, particularly for cases where there is little radiometric difference between the FOVs (due to homogeneous clouds) or when the radiometric differences are due to inhomogeneous surface characteristics under clear-sky conditions. *Cloud-clearing has been selected as the baseline approach for the OMPS-IR algorithm.*

The cloud-clearing algorithm works for *N* cloud formations using *N+1* FOVs. However, the easiest way to understand this approach is to consider 2 FOVs and a single cloud formation. Given two radiances, R, the cloud-cleared radiance for channel *i* is given by:

$$R_{i,CLR} = R_{i,1} + \eta\left(R_{i,1} - R_{i,2}\right)$$

Where η may be written in terms of the cloud fraction for each FOV (Chahine, 1974):

$$\eta = \frac{\alpha_1}{\alpha_2 - \alpha_1}$$

Note that knowledge of α is not required for the calculation of η. Also, η is spectrally independent and may be determined iteratively:

$$\eta_j = \frac{R_{i,CLR,j} - R_{i,1}}{R_{i,1} - R_{i,2}}$$

In practice a number of channels are used in the determination of η in order to reduce the magnitude of errors which results from local errors in the temperature profile and sensor noise.

Also, this technique works best for a large contrast between the FOV radiances. Rather than using all of the sensor channels, only the subset most sensitive to cloudy radiances is used (e.g. the 15 μm band CO_2 channels whose weighting functions peak above the tropopause would not be included in the cloud-cleared radiance calculation).

2.4.3 Implementation of Clouds in OMPS-IR Algorithm

This section describes the parameters required for accurate simulation and retrieval of cloudy radiances and the way that these are implemented in the OMPS-IR algorithms.

2.4.3.1 Radiance Simulation Requirements

The generation of meaningful CrIS SDRs for cloudy FOVs requires the ability to accurately simulate the spectral and spatial characteristics of clouds. Because the CrIS FOV is relatively large, the algorithm is not required to model the detailed microphysical structure of the clouds. However, the algorithm must have the ability to simulate the vertical and horizontal inhomogeneous characteristics of cloudy scenes. An additional parameter which is linked to the cloudy retrievals is the surface type – scenes with only a small contrast between the surface radiance and the cloud-top radiance are more difficult for the retrieval than cases where there is a large contrast between the surface and the cloud. In general, low clouds pose more problems than high clouds, except in the Polar regions where even cold, high clouds may exhibit little contrast with the cold surface. Thus there are four primary requirements for the accurate simulation of clouds: (1) the algorithm must have the ability to place the clouds at all levels; (2) the algorithm must allow for multiple cloud types within each FOV; (3) the algorithm should allow for a unique cloud fraction for each of the cloud types within a given FOV; (4) the algorithm should allow for a unique surface emissivity to be specified for each FOV.

2.4.3.2 Retrieval Procedure

There are several different procedures for the retrieval algorithm depending on the type of cloud mitigation approach that is selected. For cloud-clearing, the baseline OMPS-IR approach, it is necessary to compute the cloud-cleared radiances before each iteration of the retrieval. This is done by clustering the individual FOVs within the FOR to select a “clear scene” radiance. The CrIS sensor has 9 FOVs within a FOR in order to match the AMSU-A (later ATMS) footprint, and different clustering schemes may be used to determine the number of clear FOVs. Theoretically up to 8 cloud formations may be detected using the 9 FOVs. In practice it is better to reduce the number of allowed formations in order to improve the signal averaging characteristics of the retrieval. After the radiances have been clustered, the cloud-clearing algorithm is applied to the SDR radiances to create “cloud-cleared radiances”. Finally, the retrieval of atmospheric parameters is performed on the cloud-cleared radiances. If the retrieval has not converged, the cloud-cleared radiances are re-computed, and the retrieval is repeated.

The current version of the OMPS-IR algorithm uses the CrIS EDRs as the first-guess profiles. Consequently there are two options for the implementation of cloud-clearing: (1) the OMPS-IR algorithm uses the CrIS SDRs and performs the cloud-clearing using the CrIS EDRs; or (2) the OMPS-IR uses cloud-cleared radiances derived by the CrIS EDR algorithm. The advantage of the first approach is that the OMPS algorithm does not rely on the CrIS cloud-clearing implementation and FOV clustering scheme. However, it would require the CrIS algorithm to provide more specific information about which FOVs were used to derive the reported EDRs, and would

increase the computation time of the OMPS-IR algorithm. Consequently we have selected the second approach: *the baseline OMPS-IR algorithm will use the cloud-cleared radiances derived by the CrIS EDR algorithm in conjunction with the reported CrIS EDRs.* This will require the CrIS algorithm to compute and output cloud-cleared radiances for all of the CrIS channels, rather than only computing them for the channels used by the CrIS EDR algorithm. This issue is addressed in AER AIPT memo AER-AIPT-009.

2.4.4 Impact of Clouds on OMPS-IR Ozone Retrieval

2.4.4.1 Selection of Test Cases to Determine OMPS-IR Retrieval Performance

The OMPS-IR algorithm is primarily used for the retrieval of ozone under nighttime conditions. Thus it is imperative that most of the cloud simulations and analyses be performed for these types of cases. However, daytime simulations at all solar zenith angles are also important for assessing the retrieval accuracy over a variety of conditions.

Hahn et al. (1995) have analyzed 10 years of surface weather observations (1982-1991) to determine the total cloud cover and frequency of occurrence of clear sky, fog, and precipitation. This study was done worldwide for both land and ocean cases and confirms some general assumptions of the global distribution of clouds: (a) the average cloud cover is less over land than over ocean; (b) the latitudinal variation of cloud cover is greater over land than over ocean; and (c) the peak cloudiness in the intertropical convergence zone moves from 7º N in June-July-August to 2º N in December-January-February over the ocean, but to as far as 12º S over land.

Another source of cloud information is from the Atmospheric Radiation Measurement (ARM) program which operates tropical, mid-latitude, and polar observation sites. Of particular relevance to OMPS are the two different high-latitude observation stations which provide extensive information about the atmospheric and cloud properties. The primary high-latitude ARM station is located on the North Slope of Alaska, centered on Barrow. The other, temporary, station was part of the SHEBA program where an ice breaker was deployed within the perennial Arctic Ocean ice pack for 12 months starting in October 1997. The location of these sites is shown in Figure 2.4-1.

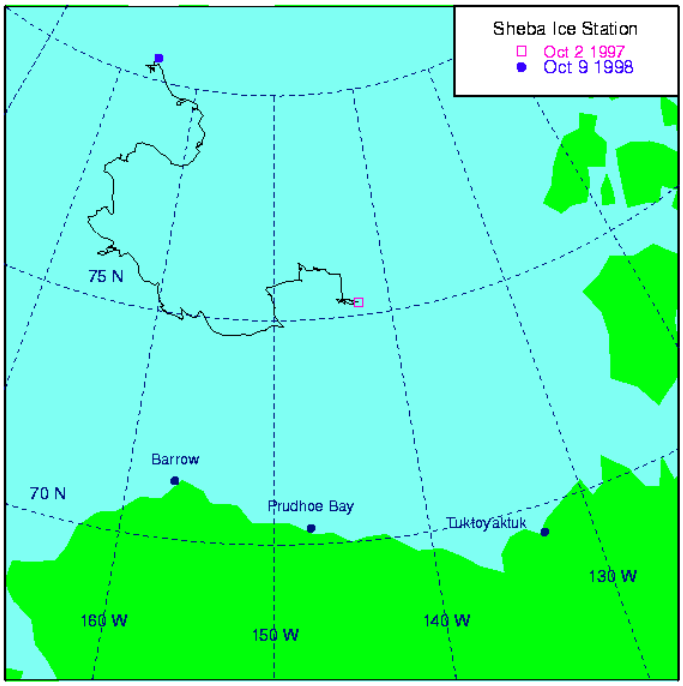


Figure 2.4-1. SHEBA ice station drift track (figure from http://sheba.apl.washington.edu).

Both of these measurement sites obtain cloud information through use of a ceilometer manufactured by Vaisala. The Vaisala ceilometer is a self-contained, ground-based, active, remote-sensing device designed to measure cloud-base height at up to three levels and potential backscatter signals by aerosols. Model CT25K, used by ARM, has a maximum vertical range of 25 km. The ceilometer transmits near-infrared pulses of light and the receiver telescope detects the light scattered back by clouds and precipitation. In addition to cloud layers it can also detect whether or not there is precipitation or other visual obstructions.

ARM data was obtained from the Atmospheric Radiation Measurement (ARM) Program website. sponsored by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Environmental Sciences Division. Cloud cover statistics were calculated for a subset of the available data covering the SHEBA measurement period and are presented in Table 2.4-1 and Table 2.4-2. The approximate pressure levels for the cloud altitudes are given in Table 2.4-3 for reference. The cloud-cover statistics are similar to those of Schweiger et al. (1999). The large standard deviation from the mean cloud altitude is indicative of the large range of cloud altitudes present. The occurrence of two or three cloud layers may be larger than that indicated in Table 2.4-1 since in many instances the lowermost cloud is optically thick to the ceilometer. This may also explain the contradiction between the occurrence of three cloud layers (typically less than 1%) and the study by Curry et al. (1996) which reported the common occurrence of five well-defined cloud layers. The large percentage of cases with total obscuration (typically rain, snow, or fog) must also be taken into account when devising both the cloud mitigation plans and the algorithm performance simulation scenarios.

Table 2.4-1. Cloud-cover occurrence statistics from ice station SHEBA and the Barrow, Alaska (*) ARM facility.

		Frequency of Occurrence				
Month	Number of Data Points	Clear Sky	1 Cloud	2 Clouds	3 Clouds	Obscured
Oct 1997	108212	13.4 %	67.6 %	5.7 %	0.21 %	13.0 %
Nov 1997	161352	28.5 %	41.9 %	7.3 %	0.47 %	21.7 %
Dec 1997	25755	36.3 %	23.8 %	1.4 %	0.12 %	38.4 %
*Dec 1997	45778	54.4 %	31.1 %	5.8 %	0.28 %	8.3 %
Jan 1998	82086	66.9 %	20.7 %	1.8 %	0.06 %	10.6 %
Feb 1998	154779	61.5 %	28.4 %	3.5 %	0.20 %	6.5 %
*Feb 1998	95564	55.2 %	35.1 %	3.5 %	0.10 %	6.1 %
Mar 1998	117567	36.3 %	46.0 %	4.2 %	0.22 %	13.3 %
Apr 1998	144660	33.5 %	43.7 %	3.2 %	0.13 %	19.5 %
*Apr 1998	35764	38.6 %	42.6 %	3.3 %	0.04 %	15.5 %
May 1998	169895	29.9 %	62.2 %	3.2 %	0.15 %	4.4 %
Jun 1998	166844	25.7 %	60.4 %	6.6 %	0.26 %	7.0 %
*Jun 1998	172065	24.2 %	59.2 %	3.8 %	0.09 %	12.7 %
Jul 1998	178499	17.1 %	54.5 %	6.4 %	0.40 %	21.6 %
Aug 1998	172049	3.3 %	67.1 %	13.5 %	0.94 %	15.2 %
*Aug 1998	112724	5.7 %	73.5%	11.7 %	0.6 %	8.5 %
Sept 1998	172505	4.2 %	70.1 %	13.7 %	1.09 %	10.8 %
*Oct 1998	113604	21.6 %	61.5 %	7.3 %	0.44 %	9.1 %

Table 2.4-2. Cloud altitude statistics from ice station SHEBA and the Barrow, Alaska (*) ARM facility.

	Cloud Altitude Mean / Standard Deviation (meters)		
Month	Single Cloud	Second Cloud	Third Cloud
Oct 1997	594 / 570	1047 / 547	1322 / 532
Nov 1997	806 / 843	915 / 677	985 / 565
Dec 1997	1485 / 1013	1968 / 587	1782 / 333
*Dec 1997	1470 / 713	1545 / 669	1801 / 562
Jan 1998	541 / 821	792 / 494	946 / 326
Feb 1998	1139 / 1001	1065 / 819	1119 / 672
*Feb 1998	1018 / 890	896 / 809	1162 / 755
Mar 1998	919 / 719	1185 / 594	1198 / 491
Apr 1998	876 / 795	1143 / 618	1188 / 465
*Apr 1998	745 / 816	764 / 545	986 / 522
May 1998	488 / 526	719 / 384	618 / 233
Jun 1998	589 / 858	783 / 494	716 / 268
*Jun 1998	1029 / 1205	1215 / 903	945 / 605
Jul 1998	992 / 1288	890 / 728	692 / 357
Aug 1998	522 / 810	811 / 574	717 / 359
*Aug 1998	764 / 1023	1032 / 780	908 / 550
Sept 1998	858 / 951	1133 / 772	931 / 534
*Oct 1998	988 / 723	1229 / 806	1330 / 738

Table 2.4-3. Standard atmosphere pressure values corresponding to 1 km levels of the atmosphere (from Anderson et al., 1986).

Altitude	Sub-Arctic Summer	Sub-Arctic Winter
0.0 km	1010 mb	1013 mb
1.0 km	896 mb	888 mb
2.0 km	793 mb	778 mb
3.0 km	700 mb	680 mb
4.0 km	616 mb	593 mb
5.0 km	541 mb	516 mb

Another way to view the ARM cloud data is to translate the temporal spacing of the point measurements into a spatial distribution of cloud types. That is, determine what types of patterns can be seen in the zenith measurements that could be used to form realistic scenes for the CrIS field-of-regard. Two examples of the spatial distribution in Figure 2.4-2 and Figure 2.4-3, for Barrow and SHEBA, respectively. The Barrow data for April 1998 shows a persistent layer of low clouds (lowest 200 m) followed by a descending range of higher clouds. In contrast, the SHEBA data for September 1998 illustrates a clear separation between two levels of clouds, all of which are found at relatively low levels.

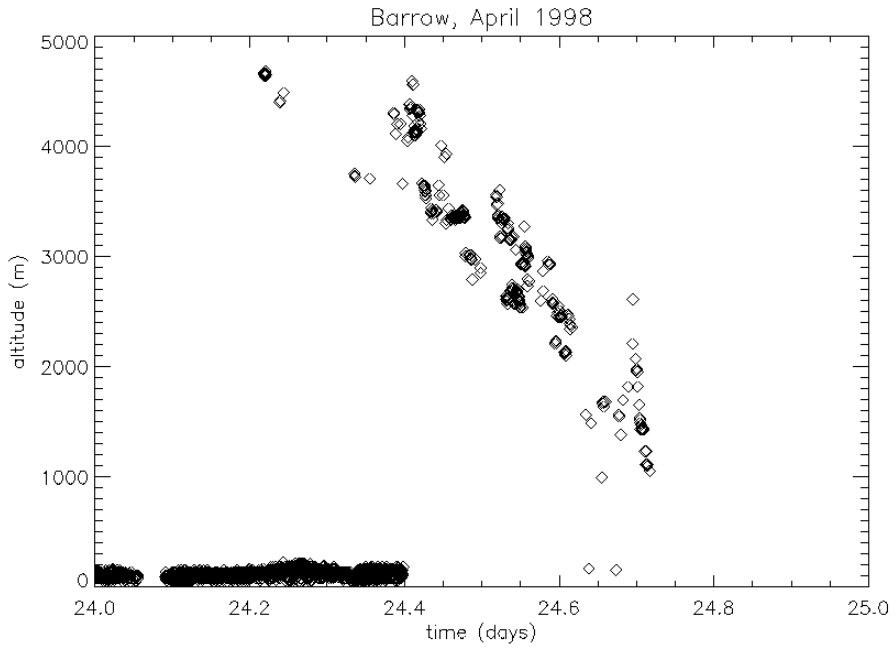


Figure 2.4-2. Altitude of lower-most cloud layer measured at the Barrow ARM site in April 1998.

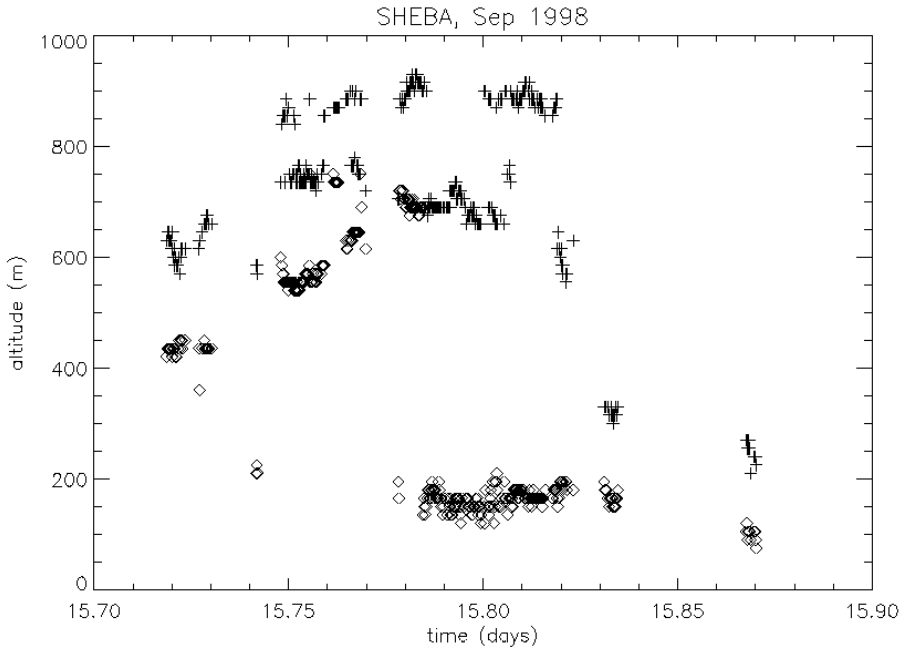


Figure 2.4-3. Altitude of lower-most cloud layer (diamonds) and next-lowest cloud (+) measured at the SHEBA site in September 1998.

The following to be completed after RRR3:
Given these statistics and patterns, what types of clouds will simulated?

2.4.4.2 Radiance Simulation Results

The following to be completed after RRR3:

Describe general results from cloud simulations

Specific retrieval performance runs belong in Section 7

2.5 Error Sources

The list of error sources that will impact the retrieval of ozone is given in Table 2.5-1. These errors can be classified as either sensor errors or algorithm errors. Details of their impact on the retrieval accuracy and precision are given in Section 7.

Table 2.5-1. Sources of Error that Impact the Retrieval of Total Column Ozone

Class	Error	Type	EDR Impact
Sensor	Radiometric Noise	Random	Precision
Sensor	Spectral Resolution	Bias	Accuracy
Sensor	Radiometric Calibration	Bias	Accuracy
Algorithm	O_3 Band Strength	Bias	Accuracy
Algorithm	Inversion Calculation	Random	Precision
Algorithm	Errors in Atmospheric Covariance Matrix	Bias	Accuracy
Algorithm	Error in T, H_2O profiles	Random + Bias	Accuracy + Precision
Algorithm	Errors in Treatment of Clouds	Random + Bias	Accuracy + Precision
Algorithm	Errors in Treatment of Surface Parameters (emissivity and temperature)	Bias	Accuracy

2.6 Principal Component Analysis of Retrieval State Vector

2.6.1 Introduction

One of the potential problems encountered in the extraction of geophysical information from remote sensing data is retrieval instability due to ill-conditioning of the background covariance matrix. This instability can be severe depending on the quality of the data used to compute the covariance matrix, the total number of retrieved parameters relative to the information content of the measurement, and the amount of inter-level correlation inherent in the retrieved geophysical parameter and included in the covariance. Thus it is always advantageous to reduce the number of retrieved parameters. Further, reducing the number of parameters also increases the speed of the retrieval process, a key element in the design of an operational retrieval algorithm.

The technique we have selected to reduce the retrieval state vector is known as Principal Component Analysis (PCA). The PCA technique allows for the representation of the geophysical profile as a linear combination of empirical orthogonal functions (EOFs). The algorithm thus performs the retrieval of projection coefficients (eigenvalues) onto the selected EOFs, rather than retrieval of the full geophysical profile. This technique has been selected by the CrIS team to reduce the size of the temperature and water vapor profile retrieval vectors from 40 level values for both temperature and water vapor to 20 eigenvalues for temperature and 15 for water vapor

(CrIS ATBD, 2000). The methodology for selecting the number of EOFs for the ozone retrieval is discussed in the following sections.

The principal component analysis is accomplished using the method of singular value decomposition (SVD). The SVD technique is based on the following theorem from linear algebra (Press et al., 1986): Any matrix **A** with *M* rows and *N* columns can be written as the product of an *M x N* column-orthogonal matrix **U**, an *N x N* diagonal matrix **w**, and the transpose of an *N x N* orthogonal matrix **V**. That is,

$$\mathbf{A} = \mathbf{U}\mathbf{w}\mathbf{V}^{T}$$

The eigenvalues are contained in the elements of **w**, which are arranged in descending order of magnitude. The EOF (**U** and **V**) matrices have the following relation:

$$\mathbf{U}^{T}\mathbf{U} = \mathbf{V}^{T}\mathbf{V} = 1$$

Thus in a geophysical sense, given the transformation matrices **U** and **V**, the appropriate **w** can be found to reproduce the truth profile matrix **A**.

The following sections discuss the application of PCA to the ozone profiles themselves and to the implementation of PCA within the OMPS-IR algorithm.

2.6.2 Analysis of Ozone Profile Database

The SVD of the ozone profile database yields a set of eigenvalues along with the appropriate EOFs. The eigenvalues are arranged in decreasing order, as shown in Figure 2.6-1. The first 8 EOFs and associated eigenvalues are shown in Figure 2.6-2. This figure illustrates that the first EOF represents the mean ozone profile while subsequent EOFs exhibit more detailed structure. Also note that the first eigenvalue is the largest, and thus has the most significant contribution to the reconstruction of the profiles, while subsequent eigenvalues decrease in magnitude.

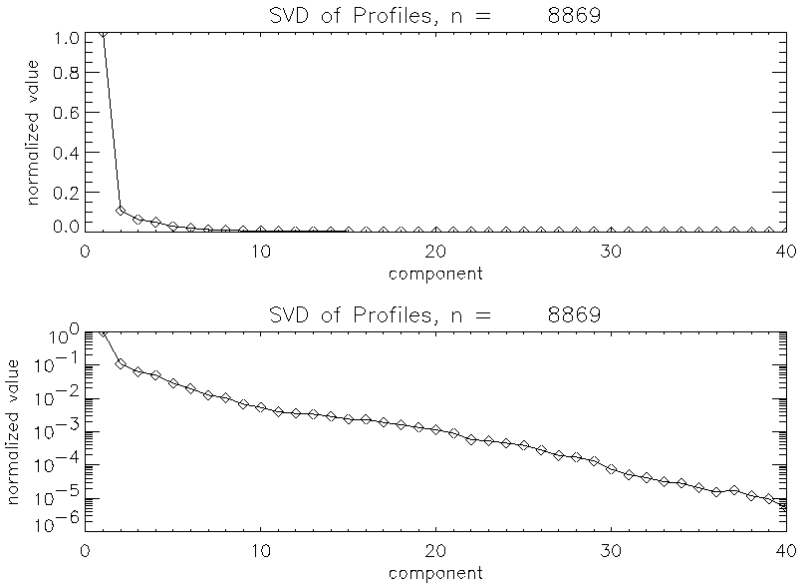


Figure 2.6-1. Eigenvalues computed from the SVD of the full OMPS-IR test database (NOAA-88, AER-Polar, and AER-POAM profiles).

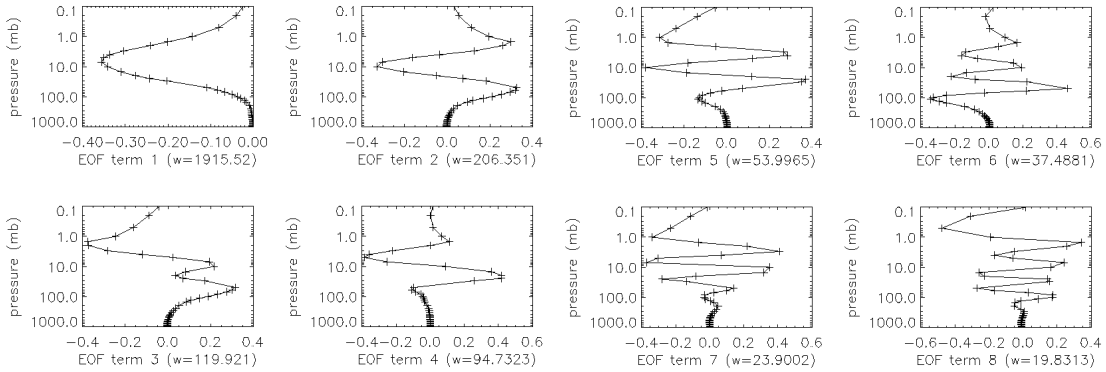


Figure 2.6-2. First 8 EOFs corresponding to the eigenvalues shown in Figure 2.6-1. The number in the lower right of each box is the non-normalized eigenvalue.

The input profiles can be reconstructed almost exactly from the EOFs, with errors limited to computer machine precision. However if some of the eigenvalues are eliminated (set to zero) prior to the reconstruction, some of the fine structure of the original profile will be lost. This is known as "representation error". Figure 2.6-3 and Figure 2.6-4 illustrate the magnitude of representation error for two types of ozone profile shape. In Figure 2.6-3 the representation error is small because there is little structure in the ozone profile, and it is easily represented by only a few EOFs. However, Figure 2.6-4 shows an ozone profile characteristic of ozone hole conditions where it is clear that more EOFs are required to accurately capture the shape of the profile.

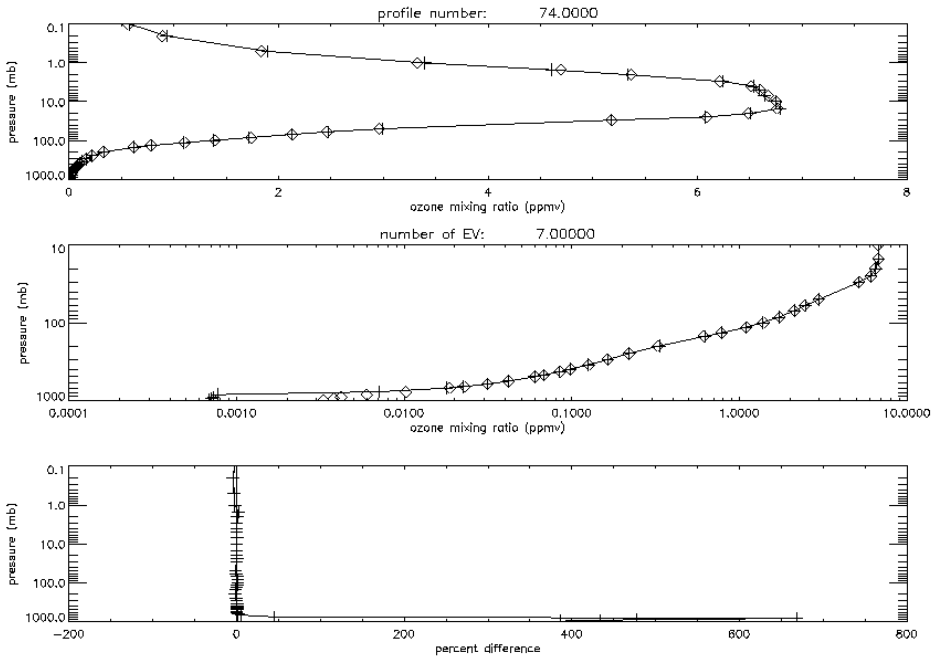


Figure 2.6-3. Comparison of original (+) and regenerated (diamond) ozone profiles when only 7 eigenvalues are used in the regeneration. The total column ozone for this profile is 384 DU and the representation error is 0.5 DU. The top and middle panels show the ozone mixing ratio in linear and log scales, while the bottom panel gives the percent error for each level.

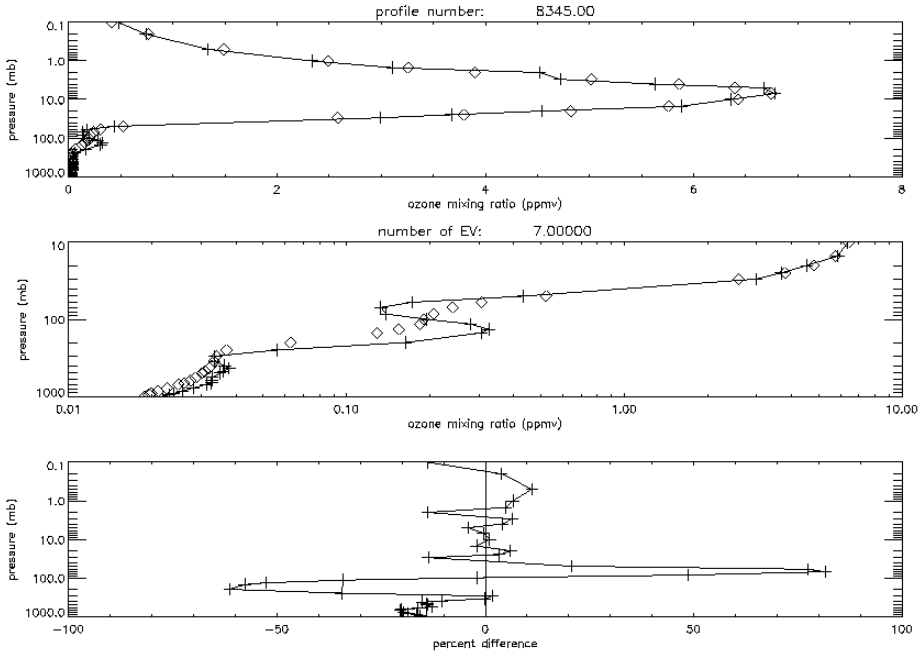


Figure 2.6-4. Same as Figure 2.6-3 except the total column ozone for this profile is 180 DU and the representation error is 10 DU.

In addition to the evaluation of the representation error for individual profiles, the overall representation errors were computed for all of the profiles in our test dataset (these are given in Table 2.6-1). Given this information it is clear that 1 EV is insufficient to represent the ozone profiles to within the EDR error threshold of 15 DU. Further, even 7 EOFs will have trouble recreating some of the profiles corresponding to large column amounts. However, the number of EOFs appropriate for the retrieval is also dependant on the information content of the measurement, as discussed in the next section.

Table 2.6-1. Representation errors (DU) for the reconstruction of the test profiles using a limited number of eigenvectors. Each of the bins is ±25 DU, centered on the named bin.

Column Ozone Bin	Number of Profiles	1 EV		7 EV		21 EV	
		Mean	S.D.	Mean	S.D.	Mean	S.D.
< 175 DU	447	47.03	1.15	5.94	0.21	1.19	0.04
225 DU	984	73.02	0.83	3.67	0.10	1.18	0.03
275 DU	3676	47.85	0.4	4.04	0.07	0.79	0.02
325 DU	2386	42.95	0.57	5.24	0.08	0.97	0.02
375 DU	1015	87.32	0.99	6.51	0.22	1.39	0.04
425 DU	251	135.08	2.06	13.55	0.77	1.32	0.09
> 475 DU	114	194.60	4.07	34.58	1.82	1.58	0.26
All	8873	58.15	0.39	5.36	0.07	1.00	0.01

2.6.3 Application to Ozone Retrieval

The application of PCA to the retrieval of geophysical parameters is relatively straightforward and is analogous to the reconstruction of the profile using only a subset of the eigenvalues. That is, if the retrieval is performed for only a subset of the total number of eigenvalues, the retrieved profile will be subject to some amount of representation error which depends on the number of eigenvalues used to regenerate the profile. This is in addition to other errors associated with the retrieval process.

There are two methods by which the PCA technique may be applied to the retrieval problem. The first method is to obtain the EOFs from a database such as NOAA-88, and use the **V** matrix in the retrieval algorithm to convert the retrieved eigenvalues (**w**) into the profile. An equivalent technique is to perform the SVD on the background profile covariance matrix:

$$\mathbf{S}_x^{-1} = \mathbf{U}\mathbf{w}\mathbf{V}^T$$

Both methods yield equivalent results, but the second approach has the advantage that the EOFs are tied directly to the covariance used by the retrieval. Since different covariance matrices are used for different classes of profiles (e.g. global versus geographic versus radiometric), the EOFs will better represent the class of profile being retrieved. Further, the SVD can be done as a part of the retrieval preprocessing, rather than as an off-line process.

As discussed in Section 2.3.2, the physical retrieval algorithm uses the following equation to solve for the state vector x:

$$x_{n+1} = x_a + \left(K^T S_y^{-1} K + S_x^{-1}\right)^{-1} K^T S_y^{-1}\left[\left(y^m - y_n\right) + K\left(x_n - x_a\right)\right]$$

where K is the Jacobian, S_y is the error covariance matrix, y^m is the measured radiance, y_n is the calculated radiance based on state vector x_n, and S_x and x_a are the background covariance and associated mean state vector. When using PCA, the eigenvector transformation results in:

$$u^T x_{n+1} = u^T x_a + \left(u^T K^T S_y^{-1} K u + u^T S_x^{-1} u\right)^{-1} u^T K^T S_y^{-1}\left[\left(y^m - y_n\right) + K\left(x_n - x_a\right)\right]$$

where u is the matrix containing the EOFs.

Determination of the appropriate number of EOFs for the OMPS-IR retrieval was performed using a set profiles that span the complete range of column ozone and skin temperature in the combined NOAA88 and Polar profile set. The profiles were divided into low, medium, and high radiance categories (computed in the 800 to 850 cm^{-1} bandpass and with divisions at 230 and 275 K in brightness temperature; see also Section 7.2.1.1.4), with 50 profiles in the low set, 111 in the medium set, and 132 in the high set. Simulated radiances were generated using the CrIS code with the baseline noise model. The final ozone retrievals were performed using CrIS-retrieved ozone and atmospheric parameters (i.e., temperature and water vapor) to specify the first guess and error covariance matrix. "Single-Noise" accuracy statistics for each EOF run were computed as a function of ozone amount in eight 50 DU bins beginning at 100 DU. "Overall" accuracy statistics were computed for each EOF run from the number-weighted average accuracy of all bins, with the results shown in Figure 2.6-5. It is clear that there is a limit to which the addition of eigenfunctions will improve the retrieval results. *Based upon this type of analysis, along with the representation error analysis described in Section 2.6.2, we have concluded that 7 eigenfunctions is optimum.*

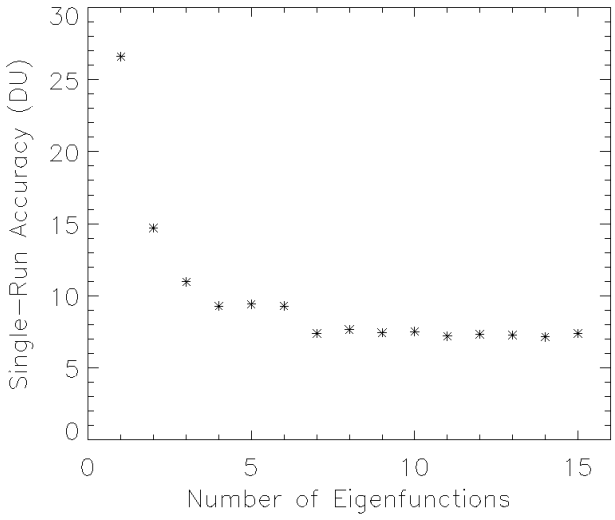


Figure 2.6-5. Total column ozone accuracy as a function of the number of ozone eigenfunctions used in the retrieval.

3 Algorithm Description

The IR total column ozone algorithm (IR-TCA) retrieves total column ozone from the CrIS SDRs using the CrIS EDRs (pressure, temperature and moisture profiles) as first-guess profiles. The optimal spectral sampling technique (OSS, described in Section 2.3.1) is used to compute the forward model radiances, while the optimal estimation approach (Section 2.3.2) is used for the inversion. A key feature of the IR-TCA is the advanced handling of clouds. The algorithm is modular and, once the final CrIS design is known, the algorithm will be constructed to accurately treat homogeneous overcast, partial cloudiness, and complex fields with optically thin cirrus.

3.1 Sensor Data Record (SDR) Production

The CrIS SDRs will be produced by the CrIS team, rather than as a part of the IR-TCA effort. Our assumptions about the CrIMSS (CrIS IR and microwave) SDR content, format, and calibration are outlined in Section 2.2. Specific requirements placed upon the CrIS by the IR-TCA are given in Section 5.1.

3.2 Environmental Data Record (EDR) Production

3.2.1 Retrieval Algorithm Description

The IR total column algorithm (IR-TCA) extracts total column ozone from the CrIS SDRs and is constructed to allow for the simultaneous retrieval of total column ozone from N fields-of-view (FOV) within a given field-of-regard (FOR). The IR-TCA is highly modular to allow easy adaptation to changes in the sensor design or upgrades to the physics included in the model. It has been coded with multi-threading capabilities in order to take advantage of machines with parallel processing architecture.

For the forward radiance model the IR-TCA uses the optimal spectral sampling (OSS) technique. OSS is an AER-developed radiative transfer model that has been extensively tested against LBLRTM (a state of art “exact” line-by-line model described by Clough and Iacono, 1995) with demonstrated accuracy to within 0.05 K in brightness temperature (see Section 7). This accuracy is more than sufficient for the needs of the retrieval algorithm.

In addition to its high accuracy relative to LBLRTM, OSS provides a means of extremely fast computation of the radiances and analytical derivatives required by the inversion module. If the 20-minute NPOESS requirement for processing time is reduced to 16 minutes to allow for SDR generation from raw data records (RDRs) and other processes, the time allotted for the retrieval algorithm becomes 0.0047 seconds per FOV (assuming 205,000 FOVs per orbit). Since the current (1999) version of OSS requires about 0.47 seconds per FOV for the entire retrieval process on current technology machines, the OSS is close to meeting the year 2007 processing requirements, but still requires some optimization. This rapid processing speed is the result of a number of novel features derived from ongoing algorithm trade studies. One such feature is the use of an empirical orthogonal function (EOF) representation of temperature, moisture, and ozone profiles.

The retrieval module for the IR-TCA utilizes the optimal estimation (OE) technique to conduct the radiance inversion (Rodgers, 1976, 1990). This method was selected because it is a robust and proven technique which allows us to accurately account for the sensor and atmospheric

characteristics of the measurements. OE is a physical-statistical inversion method that optimally incorporates both the physical radiative properties of the measurement as well as the known statistical correlations of the ozone and other atmospheric parameters. The latter can indeed be updated and made more accurate through the use of data from the OMPS UV suite. We have developed versions of this technique for a number of different programs, including EOS-TES, MSX, and CIRRIS-1A.

Together the OSS and OE algorithms provide a robust, accurate, computationally efficient means of obtaining total column ozone from the CrIS sensor. A key feature of our retrieval algorithm is the advanced handling of clouds, and the algorithm will be constructed to accurately treat homogeneous overcast, partial cloudiness, and complex fields with optically thin cirrus.

3.2.2 Assumptions

All assumptions about the SDRs and EDRs produced by the CrIS sensor and the CrIMSS suite are given in Section 2.2, with specific requirements given in Section 5.1.

3.2.3 Method

The flow of data from the CrIS SDR to the total column ozone product is shown in Figure 3.2-1. The CrIS EDRs (pressure, temperature, and water vapor profiles) and ozone climatology are used as first-guess profiles. Along with ozone, the algorithm simultaneously retrieves temperature, water vapor, surface temperature and emissivity in order to improve the retrieval result. The CrIS retrieval error covariance is used to constrain the first guess values which are obtained from CrIS EDRs. A data quality check, described in Section 3.2.4, is used to provide a measure of the retrieval accuracy.

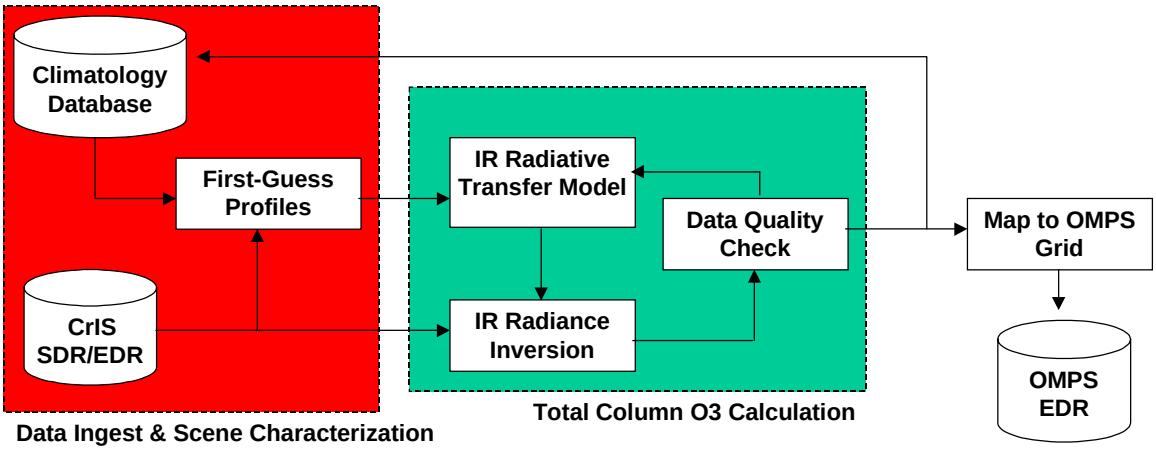


Figure 3.2-1. Overview of the Infrared Total Column Ozone Algorithm (IR-TCA).

The retrieval algorithm uses only a subset of the full range of CrIS channels. Specifically, only the channels from 950-1095 cm^{-1}. The retrieval is performed for temperature, water vapor, ozone, surface emissivity, and surface skin temperature.

3.2.4 Data Checks/Assessment

Along with the creation of the total column ozone EDR, a quality control data assessment is performed to provide an indication of the overall accuracy of the retrieval product. This data flag will be the result of an examination of radiance residuals, a determination of the consistency with nearest neighbor retrievals, and (when available) a check of consistency between the IR and UV-nadir retrievals. The IR-TCA is designed with the concept of 'graceful degradation' so that in the absence of CrIS SDRs the algorithm will revert to either climatological or nearest neighbor retrieval values. These fill data will be marked with flags to distinguish them from ozone measurements using CrIS data. Details are given in the quality control appendix (Section 12).

4 Assumptions

Table 3.2-1 summarizes the assumptions inherent in the IR-TCA.

Table 3.2-1. Assumptions of the IR Total Column Ozone Forward Model and Retrieval

Forward Model
The atmosphere can be represented by Beer's Law
The atmosphere's lower bound consists of a Lambertian reflecting surface for the ground
The sensor errors can be characterized through the use of radiometric noise and radiometric bias terms
Retrieval
The profile shape, which is critical for retrieving accurate total column ozone, can be adequately retrieved using empirical orthogonal functions (EOFs) that define the profile shape
The CrIS EDR performance is consistent with the output of the CrIS error covariance matrix
The CrIS EDR performance is not significantly worse than the EDR threshold requirements

5 Input Data Requirements

5.1 Primary Sensor Requirements

The OMPS-IR total column algorithm (IR-TCA) requires data from CrIS EDRs and SDRs. We assume that the SDRs and EDRs are as defined in Section 2.2. The key assumptions employed in the IR-TCA development are summarized as follows. The requirements for cloud-cleared radiances and the error covariance matrix were outlined in the AER AIPT memo AER-AIPT-009.

1) The radiances are the essential information used by the IR-TCA. We assume that the SDRs provide calibrated apodized radiances with a well-known instrument line shape (ILS). Further, we assume that this calibration will include corrections for the off-optical axis effects, such as spectral shift, self-apodization, phase distortion, etc., and that the channel centers will be mapped on a fixed wavenumber grid to standardize the IR-TCA forward model sensor parameters.
2) The SDRs contain information about the radiometric noise, radiometric uncertainties, spectral shift errors, band-to-band co-registration errors, LOS jitter errors, etc. which are

essential to define the sensor error covariance matrix used in the IR-TCA inversion. We do not anticipate that the actual values will necessarily be passed for each FOV, but only that there is some provision for providing updates if these values change.

3) The SDRs provide the CrIS cross-track scan angle for determining the appropriate atmospheric path geometry.
4) The SDRs provide geolocation information (latitude and longitude) which can be used for determining the appropriate first-guess ozone profile and surface emissivity.
5) The CrIS SDRs and EDRs contain data quality flags. These flags will be used to estimate the quality of the SDRs and EDRs to determine whether or not the IR-TCA should use a particular SDR pixel or retrieval EDR.
6) Channels in band 1 contain 950-1095 cm^{-1}.
7) Cloud-cleared radiances produced by the CrIS EDR retrieval code, including the 950 – 1095 cm^{-1} region, are provided as part of the EDR output.
8) The error covariance matrix from the CrIS EDR retrieval, $S = \left(S_x^{-1} + K^T S_\varepsilon^{-1} K\right)^{-1}$, is provided as part of the EDR output.

The IR-TCA is designed to utilize channels from all three bands of the CrIS sensor. Once the final CrIS design and noise characteristics have been determined, a channel selection may be performed in order to minimize the actual number of channels used in the retrieval. The retrieval problem is such that in a graceful degradation mode the algorithm will work as long as there are data available from Band 1 of CrIS (nominally 650 – 1095 cm^{-1}).

5.2 Other OMPS Sensor Data Requirements

The IR-TCA does not require additional OMPS sensor data. However, data from the OMPS nadir sensor could be used to provide updated first-guess total column ozone information for solar zenith angles less than 80 degrees. Similarly the OMPS limb sensor could provide ozone profile information that would be used to improve the first-guess ozone profile shape. These items have been postponed for further study.

5.3 Other NPOESS Sensor Data Requirements

Clear sky: **None**
Cloudy sky: **TBD**

5.4 Climatology Data Requirements

The IR-TCA requires two forms of climatological data, “static” data and “transient” data. Static data consists of information that is provided once and may or may not be updated as better information becomes available, while transient data is updated on a regular basis. Essential static information is a digital elevation map for defining the lower boundary, and a land/ocean surface database to set up the emissivity first-guess. Both of these requirements will be satisfied with the 1 km resolution topographic map provided by the USGS (see link reference under USGS). Essential transient data include an updated ozone climatology from the OMPS nadir sensor, the IR algorithm itself, or other active satellite and ground station ozone sensors, and updated ozone profile shape climatology based on the OMPS limb sensor. Because this ozone climatology will be updated on a regular (TBD) basis, it will include the impact of seasonal changes as well as latitudinal and longitudinal dependencies.

6 Output Data Description

6.1 EDRs

The main data product is the total column ozone in Dobson Units.

6.2 EDR Content

The content of the IR Total Column Ozone output data records shall include the following:

Header record with:

- EDR identification (e.g. “total column ozone”)
- Spacecraft identification
- Sensor identification
- Date and time of EDR generation
- Date and start time of data in file
- Date and end time of data in file

Data records (approximately 233,100 IFOV per orbit) with

- Orbit number
- Date and time of data
- Data acquisition orbit number
- Data transmission orbit number
- Ascending node Julian date and time tag
- Spacecraft altitude
- Identification of SDR calibration parameters
- EDR Algorithm identification number
- EDR Algorithm version number
- CrIS channels used in retrieval
- Latitude and longitude of scene
- Solar zenith angle
- Satellite zenith angle
- Solar azimuth angle (the angle between the sun and the satellite IFOV)
- Total Ozone (EDR)
- Surface category(land/ocean flag)
- Data quality flags
- Additional data products (see Section 6.3)

6.3 Additional Data Products

This section to be completed in detail after RRR3

Data quality flags
Ozone profile (?)
Error covariance matrix (?)
Cloud-clearing parameter (?)

7 System Accuracy and Precision

This section describes the methods used to estimate accuracy and precision, the types of profiles used in the simulations, the magnitude of these error sources and their impact on the accuracy and precision retrieval statistics. A number of different sources of error impact the accuracy and precision of the total column ozone retrieval (see Section 2.5). These errors can be divided into two categories, errors that are attributable to the sensor (Section 7.2.2), and errors from the algorithms (Section 7.2.1). Baseline performance estimates (Section 7.3) include all sensor and algorithm error sources.

7.1 Introduction

This section provides definitions of accuracy and precision, a description of the profiles chosen for these calculations, and a description of the processes used to generate baseline retrieval performance estimates.

7.1.1 Definitions of Measurement Accuracy and Precision

The definitions for the calculation of accuracy and precision are taken from Appendix A of the NPOESS Sensor Requirements Document (17 March 1997) and are repeated in Sections 7.1.1.1 and 7.1.1.2 for convenience. The method by which we have implemented these definitions into our estimates of system accuracy and precision is described in Section 7.1.1.3.

7.1.1.1 Measurement Accuracy

Measurement accuracy is defined as the magnitude of the difference between the mean estimated value of a parameter and its true value. This estimate may be the result of a direct measurement, an indirect measurement, or an algorithmic derivation. The mean is based on a set of estimates satisfying the following two conditions:

(1) The set is large enough so that the sample size error in the measurement accuracy is much smaller than the specified measurement accuracy value.
(2) The true value of the parameter is the same for all estimates in the set.

The second condition is imposed because a measurement accuracy requirement must be met for any true value of the parameter within the measurement range, not in an average sense over the measurement range. In practice, such as in the analysis of simulation results or measured calibration/validation data, it is understood that measurements will be binned into sets for which the true value of the parameters falls into a narrow range, preferably a range much smaller than the required measurement range.

For an ensemble of N estimates of the parameter x, the measurement accuracy β_N is given by the following formula:

$$\beta_N = \left|\mu_N - x_T\right| \tag{44}$$

where: μ_N is the sample mean, x_T is the true value of the parameter, and |…| denotes the absolute value. The sample mean μ_N is given by the following formula:

$$\mu_N = \frac{\sum_{i=1,N} x_i}{N} \tag{45}$$

where x_i is the value obtained in the ith estimate of the parameter x and $\Sigma_{i=1,N}$ denotes summation from i=1 to i=N.

7.1.1.2 Measurement Precision

Measurement precision is defined as the standard deviation (one sigma) of an estimated parameter. This estimate may be the result of a direct measurement, an indirect measurement, or an algorithmic derivation. The standard deviation is based on a set of estimates satisfying the following two conditions:

(1) The set is large enough so that the sample size error in the measurement precision is much smaller than the specified measurement precision value.
(2) The true value of the parameter is the same for all estimates in the set.

The second condition is imposed because a measurement precision requirement must be met for any true value of the parameter within the measurement range, not in an average sense over the measurement range. In practice, such as in the analysis of simulation results or measured calibration/validation data, it is understood that measurements will be binned into sets for which the true value of the parameters falls into a narrow range, preferably a range much smaller than the required measurement range.

For an ensemble of N estimates of the parameter x, the measurement precision σ_N is given by the following formula:

$$\sigma_N = \left[\frac{\sum_{i=1,N} (x_i - \mu_N)^2}{(N-1)} \right]^{1/2} \tag{46}$$

where: μ_N is the sample mean (defined in the definition of measurement accuracy), x_i is the value obtained in the ith estimate of the parameter x, and $\Sigma_{i=1,N}$ denotes summation from i=1 to i=N.

7.1.1.3 Method Employed for Accuracy and Precision Calculations

The impact of sensor and algorithm errors on the OMPS IR total column ozone retrievals is simulated using the current baseline CrIS sensor specifications (given in Section 2.2). An important part of the sensor and algorithm error studies is the ability of the simulations to capture a variety of conditions representative of the global variability of the atmosphere. To accomplish this, atmospheric profiles ranging from polar to tropical, winter to summer, and land/ocean conditions were selected from the NOAA-88 atmospheric database. In addition, we have supplemented this data set by using SAGE-2 data, ozonesonde data, and model outputs to

construct profiles representative of ozone hole conditions. These data sets are described in Section 7.1.2. Individual test results are presented in the form of scatter plots of retrieval accuracy/precision versus the total column ozone amount for each of the profiles included in the test, while total system performance is given with bar graphs for ranges of column ozone (Section 7.3).

Another key aspect to the calculation of accuracy and precision is the number of simulations done for each profile tested, i.e. the value of "N" in Equation 45 and Equation 46. If the value of N is too low, the simulations will not present an accurate picture of expected results. However, if N is too high then the simulations will take more time to complete than is necessary for obtaining reasonable results (an analysis of algorithm timing is given in Section 8.6). A quick test was performed on mid-latitude and polar profiles to determine a reasonable number of simulations, and the results are illustrated in Figure 7.1-1 and Figure 7.1-2. From these tests we decided that 80 simulations will provide reasonable results.

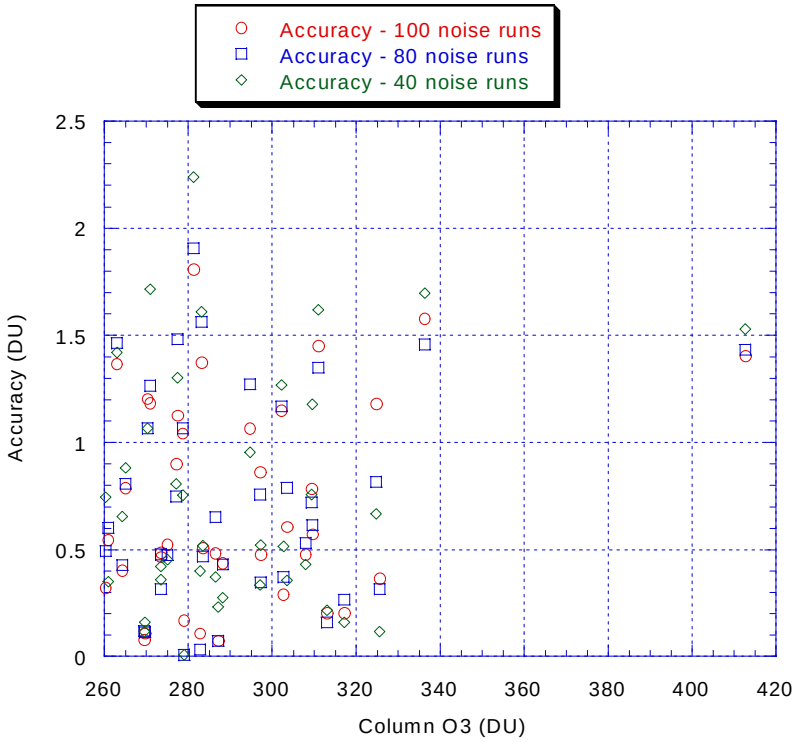


Figure 7.1-1. Accuracy calculation for mid-latitude profile retrievals with different numbers of random noise runs.

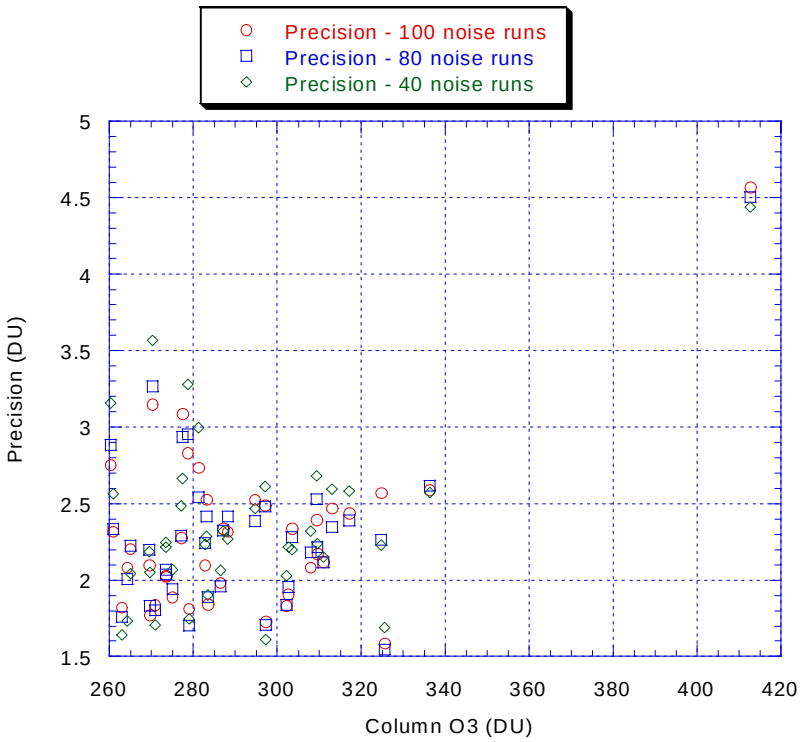


Figure 7.1-2. Precision calculation for mid-latitude profile retrievals with different numbers of random noise runs.

7.1.2 Atmospheric Profiles for OMPS-IR Algorithm Testing

A thorough understanding of the OMPS-IR accuracy and precision requires the development of a robust series of tests with which to evaluate the retrieval performance. In turn, these tests must rely on a database of profiles which encompass the range of global variability for each of the factors influencing the retrieval of total column ozone. The NOAA-88 set of profiles were provided by the IPO for testing the IR algorithms. This database consists of 8344 profiles of temperature, water vapor, and ozone, along with information about clouds and the surface characteristics. While this data is global in nature, it does not accurately reflect all of the conditions likely to be encountered. A clear example of this is the lack of profiles characteristic of ozone-hole conditions. For the test and evaluation of the OMPS-IR algorithm we have corrected this deficiency with the addition of a set of ~120 polar ozone profiles. These profiles were created by merging SAGE-2 and ozonesonde profiles with those derived from an atmospheric chemistry and transport model. The creation of these profiles is described in more detail in the AIPT memos AER-AIPT-002 and AER-AIPT-004. An additional 409 profiles were created using POAM data coupled with coincident ozonesondes. These profiles are representative of Polar conditions in general, rather than just ozone-hole conditions. The combined database of NOAA-88 and polar ozone-hole profiles will be referred to as the OMPS-IR global test database and will be supplemented (as needed) with additional profiles to test specific environmental conditions.

Algorithm run-time restrictions prohibit the continual test and evaluation using all of OMPS-IR global database profiles. Further, some of these profiles were needed to construct the atmospheric covariance matrices, making them unavailable for retrieval testing (these profiles are referred to as the 'dependent set'). Two subsets of the profile databases are needed to assure thorough test and evaluation: (1) a dependent set for use in the generation of covariance matrices, and (2) a small (360 profiles) independent subset for day-to-day testing of algorithm enhancements. A larger independent set may be reserved for additional algorithm performance estimates. These sets were divided geographically in order to allow focussed tests on a particular region (e.g. Polar cases). The atmospheric covariance matrix and the associated background/first-guess profiles used for the baseline retrieval results were derived from the global dependent set, rather than from the geographic partitions of this set. Ongoing tests will determine the applicability of a global covariance versus a covariance calculated for a particular subset of the data. Preliminary results are described in AER AIPT memo AER-AIPT-008. The total column ozone as a function of latitude for the dependent and independent profile sets is given in Figure 7.1-3. Additional profiles with ozone-hole characteristics (derived from South Pole ozonesonde data and used to supplement the NOAA-88 profiles) are shown in Figure 7.1-4. More specific information about the profiles used in the baseline performance tests is given in Section 7.3.

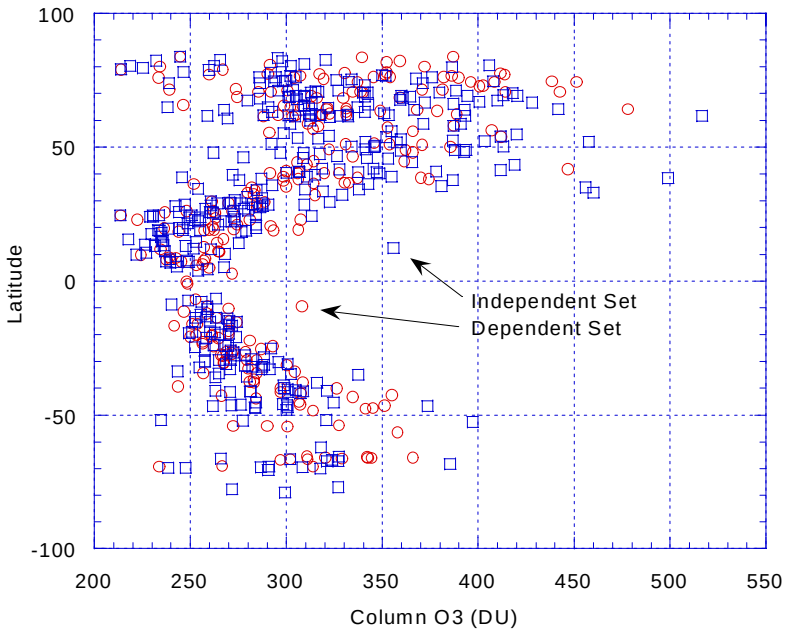


Figure 7.1-3. Comparison of the dependent and independent profiles chosen from NOAA-88.

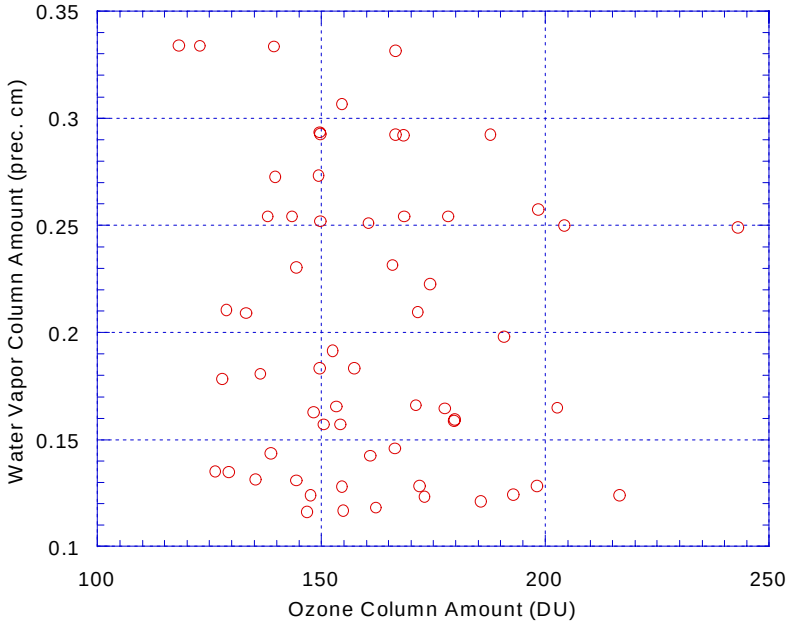


Figure 7.1-4. Polar profiles used to supplement the NOAA-88 data set.

7.1.3 Process for Generation of Performance Estimates

This section describes the process used to generate test scenes for the calculation of retrieval performance estimates…..

7.2 Sensor/Algorithm Contributions to Accuracy and Precision

7.2.1 Algorithm Impacts to Ozone Retrieval Accuracy and Precision

7.2.1.1 Algorithm Impacts to Ozone Retrieval Accuracy

7.2.1.1.1 Algorithm Accuracy: Forward Radiance Model

Because the OSS forward radiance model represents an approximation to the true line-by-line calculation, validation of the OSS is the first step in assessing the impact of different error sources. This validation must be done to ensure that the statistics generated though simulations will accurately reflect the sensor performance in an operational environment. The line-by-line radiative transfer model chosen as the reference standard is the LBLRTM model, which employs the FASCODE algorithm and has been extensively validated against atmospheric measurements [Clough *et al.*, 1992; Snell *et al.*, 1995]. Given 'perfect' input parameters, the numerical

accuracy of the LBLRTM radiance and transmittance calculations is estimated to be 0.5% (S.A. Clough, private communication). The accuracy of OSS for CrIS band 1 is critical to our ability to determine OMPS-IR performance estimates. Incorrect behavior, e.g. changes in optical depth with temperature and pressure, could lead to erroneous conclusions and improperly drive the algorithm design. Figure 7.2-1 shows that for CrIS band 1 the OSS model agrees on average to within 0.05 K of the LBLRTM brightness temperatures. These statistics were computed from a set of 100 profiles, each of which had a random viewing angle ranging from 0 to 60 degrees.

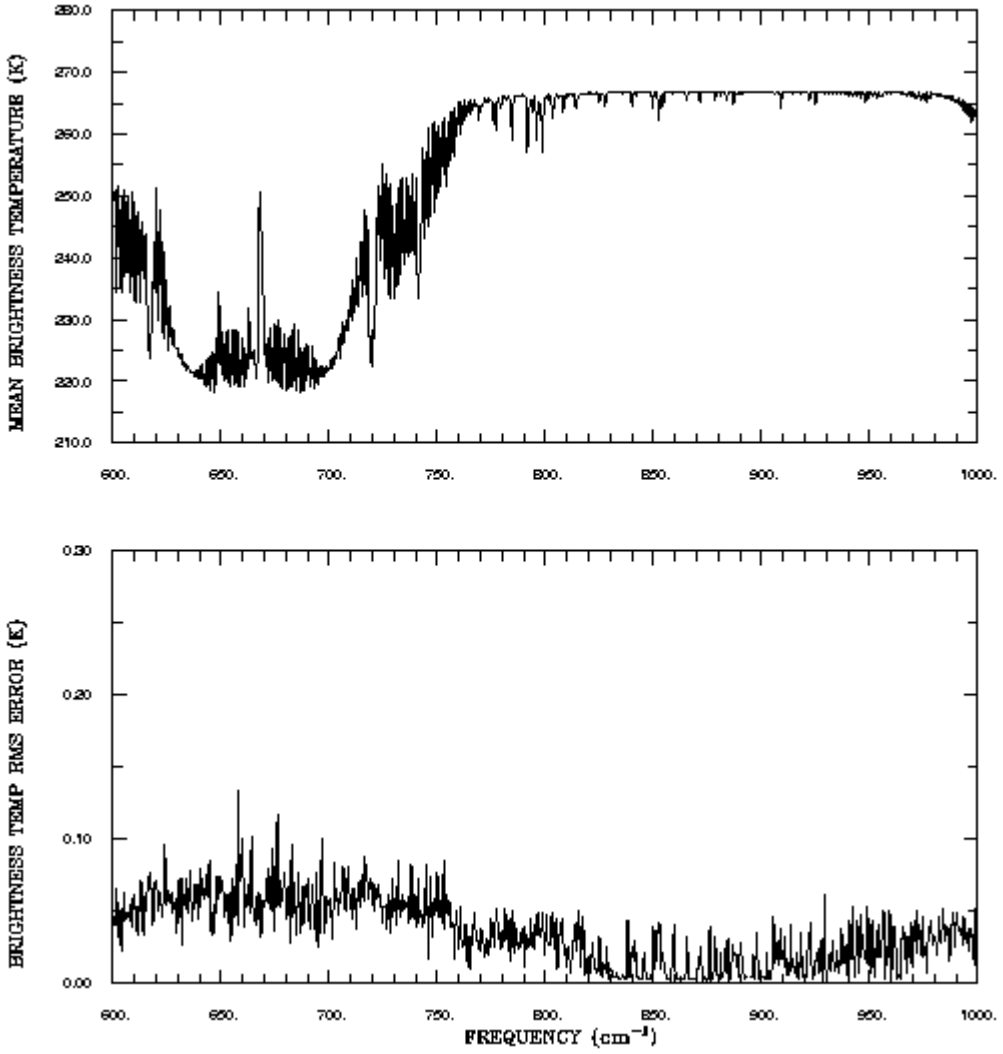


Figure 7.2-1. Comparison of OSS and LBLRTM for Band 1 of CrIS.

7.2.1.1.2 Algorithm Accuracy: Ozone Band Strength

The largest algorithm impact on the retrieval accuracy is due to the uncertainty of the ozone band strength. The infrared ozone spectral lines are typically measured in a laboratory using a Michelson interferometer. This results in a set of spectral parameters where the line shape and relative line strengths are well identified, but there is error in the overall magnitude of the entire band. The estimated band strength error of the 9.6 μm ozone band is approximately 5% (J. Russell, private communication). Multiplying the tabulated line strengths by 1.05 before

simulating the measured radiances is one way to accurately represent the nature of this error. Figure 7.2-2 illustrates the impact of ozone band strength uncertainties on the retrieval accuracy. It is clear that in many cases the estimated band errors of 5% lead to a retrieval accuracy that is greater than the 15 DU threshold requirement. A detailed post-launch calibration/validation program is expected to decrease the band errors to about 2%, greatly improving the retrieval accuracy. Note that the low ozone values (i.e. less than 210 DU) shown in Figure 7.2-2 and Figure 7.2-3 have abnormally high errors. This is due to the use of an inappropriate atmospheric covariance matrix (see Section 7.1.2). However, it is clear that the impact of reducing the spectroscopic errors is similar as for the rest of the profiles.

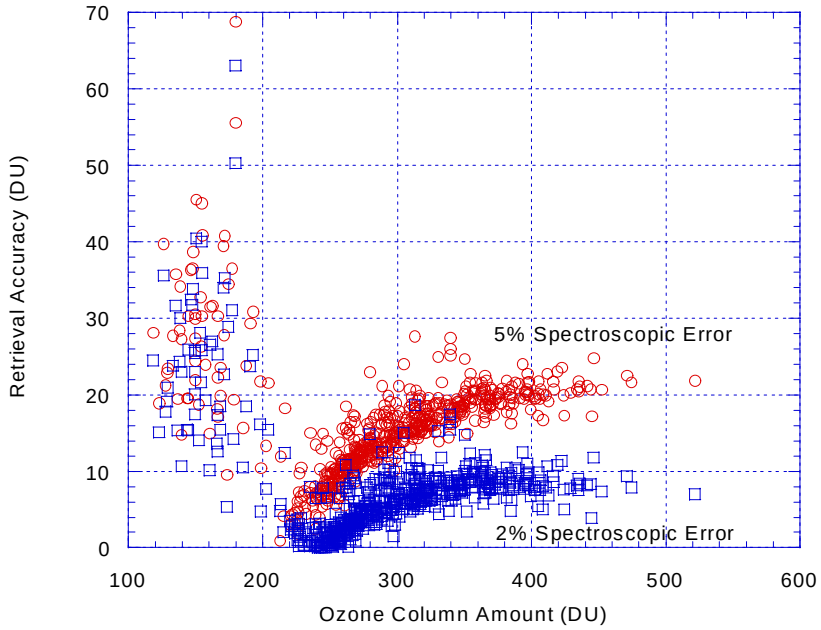


Figure 7.2-2. Impact of ozone band strength errors on the total column ozone retrieval accuracy. The threshold limit for accuracy of the ozone retrieval is 15 DU. The ozone column amounts below 210 DU represent profiles from the enhanced Polar data set.

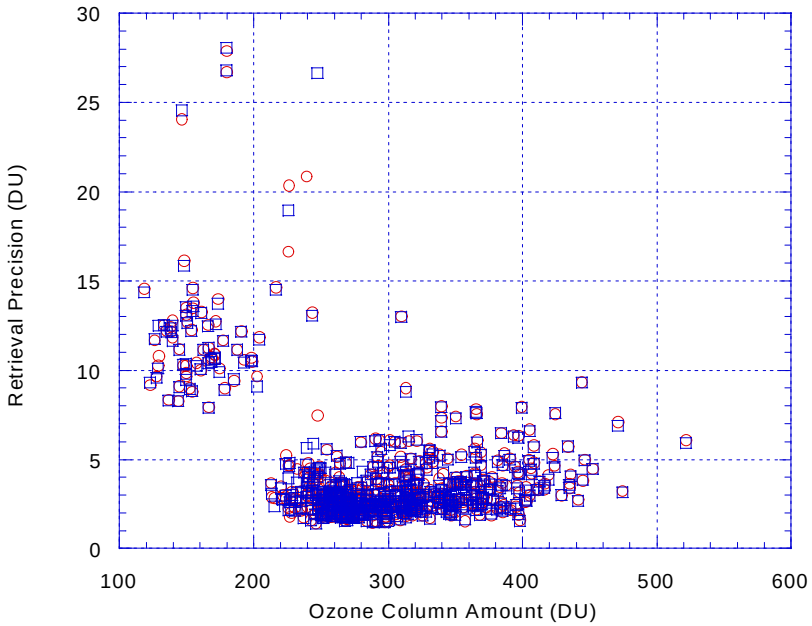


Figure 7.2-3. Impact of ozone band strength errors on the total column ozone retrieval precision. The ozone column amounts below 210 DU represent profiles from the enhanced Polar data set.

7.2.1.1.3 Algorithm Accuracy: Errors in Auxiliary Parameters

Errors in auxiliary atmospheric variables (e.g., temperature and moisture profiles, surface emissivity and surface skin temperature) also affect the accuracy of the desired retrieval quantities (the impact of clouds is discussed separately below). In the concurrent retrieval approach, where all of these variables are retrieved simultaneously with ozone, errors within given variables tend to be minimized as they are not propagated through the retrieval process. As discussed in Section 6.2, the magnitude of the radiance residuals will provide an indication of the quality of the retrieval, while the spectral location of these errors can provide insight into the source of the error. For example, if the dominant error is limited to the window channels, then the surface parameters will most likely be the dominant source of the errors.

The impact of errors in non-ozone parameters has been evaluated in the context of a 2-stage retrieval algorithm. In one such scenario the CrIS EDRs are used as "truth" and the OMPS-IR algorithm retrieves only ozone. For that case the ozone retrieval errors are unacceptably large. Further details are given in the AER AIPT memo AER-AIPT-007. However, the retrieval results are greatly improved if the CrIS EDRs are used as a first-guess for a simultaneous retrieval of all parameters. (As noted in Section 7.2.1.1.4, the error covariance must be used in place of the climatological covariance for the CrIS parameters used in the retrieval.)

7.2.1.1.4 Algorithm Accuracy: Atmospheric Covariance Matrix

The atmospheric covariance matrix is used to constrain the retrieved parameters so that they maintain reasonable and consistent values. This matrix is constructed from numerous profiles which are chosen to be representative of the range of expected measurement conditions. The diagonal elements of this matrix represent the expected variability of the retrieved quantity (e.g. the temperature at each level), while the off-diagonal elements represent the correlation between parameters (e.g. the correlation of 1000 mb and 950 mb temperatures, or the correlation between the temperature and ozone for a given level). To the extent that the profiles used to construct the covariance matrix are not representative of the measurement conditions, errors will be introduced into the retrieval. Thus it is important that the data set used to calculate the covariance matrix is able to capture the global and seasonal variability of total column ozone, temperature and moisture profiles, and surface parameters. This was the rationale for the addition of more stressing Polar cases to our profile dataset (Section 7.1.2).

Over 2000 profile sets from the NOAA-88 database were used to generate the global error covariance statistics used by the OMPS-IR retrieval algorithm in the initial stages of this project. Unfortunately, the NOAA-88 data do not adequately represent the global variability of ozone since they were based on incomplete vertical ozone measurements themselves (Larry Flynn, private communication). This led to one of the critical steps of generating accurate and statistically complete ozone and atmospheric profiles and surface parameters for use as a calibration/validation database (see Section 7.1.2). A subset of this new database was used (along with the NOAA-88 profiles) to calculate regional covariance matrices (e.g., a mid-latitude summer covariance). However, it was found that there were still many cases for which the retrieval performance did not meet the performance requirements. Further investigation showed that the poor results were primarily for cases where the profile shape (of temperature, water vapor, or ozone) was at the edge of the variance for the regional mean profile (e.g. a warm profile for the Polar set). A new approach was investigated whereby the window-channel radiances were evaluated to classify the profile type radiometrically rather than geographically or seasonally. Because the fundamental unit upon which the retrieval is based is radiance, the results are much improved using this technique. Figure 7.2-4 illustrates the partitioning of the profile data into three radiometric regimes, with the dashed lines indicating the surface skin temperature corresponding to the window radiance (240 and 275 K). Additional tests are underway to determine if three radiometric partitions are sufficient.

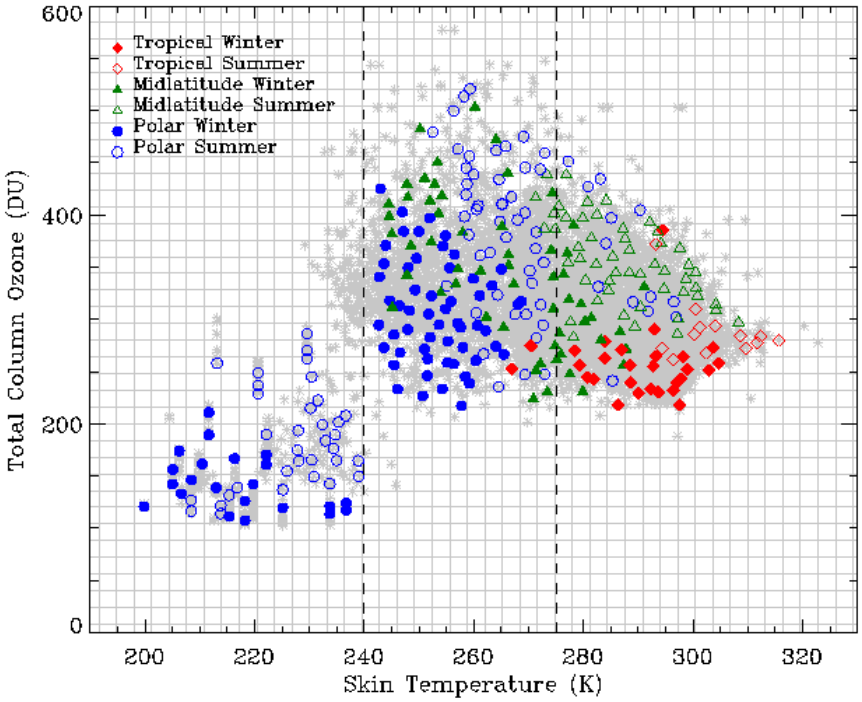


Figure 7.2-4. Division of profiles into three radiometric regimes, as indicated by the dashed lines. The points indicate that profiles which are similar radiometrically may be very different geographically.

For the case of the 2-stage algorithm (using CrIS EDRs as a first guess for the ozone retrieval, described in Section 3.2.3) the error covariance of the CrIS retrieval must be used instead of the climatological covariance. This covariance gives the expected uncertainty of the retrieval, and is given by:

$$S = \left(S_x^{-1} + K^T S_\varepsilon^{-1} K\right)^{-1}$$

In the case of no sensitivity to a parameter, or very large noise values, S will converge to the climatological covariance (indicating that the CrIS retrieval is no better than a climatological guess).

7.2.1.1.5 Algorithm Accuracy: Clouds

This section to be updated after RRR3

Clouds strongly affect the measured IR brightness temperatures through absorption and multiple scattering. Semi-transparent clouds change the spectral characteristics and magnitudes of radiances. Optically thick clouds completely absorb the IR emission, and an IR sensor essentially measures the cloud top temperature. The absorption and scattering properties are dependent upon the phase (ice versus liquid) and size of the cloud particles/drops. The spatial

distribution and inhomogeneity of clouds are also important, and the impact on the radiances depends on the sensor field-of-view. Handling the effects of clouds in the retrieval process is crucial for optimizing measurement accuracy and precision.

7.2.1.2 Algorithm Impacts to Ozone Retrieval Precision

As a test of the magnitude of the algorithm's precision, "true" profiles and surface parameters were used to calculate brightness temperatures for retrievals. Within the retrieval, the "truth" data was used as the first-guess data and there were no errors assumed in the data (i.e., no radiometric noise, sensor bias error, etc.). The final retrieved profiles were compared to the "truth" profiles, and the difference is the algorithm precision errors. The algorithm errors had a negligible effect on the overall precision of the ozone retrievals.

7.2.2 Sensor Impacts to Ozone Retrieval Accuracy and Precision

7.2.2.1 Sensor Impacts to Ozone Retrieval Accuracy

7.2.2.1.1 Sensor Accuracy: Sensor Spectral Resolution

Changes in the spectral resolution of the sensor can impact the accuracy of the total column ozone retrieval. Baseline CrIS simulations were performed prior to the OMPS preliminary design review (January 1999) using the sensor parameters given in the SRD. These results showed that a modest change in spectral resolution (e.g., changing the optical path difference by 25%) has little impact on sensor accuracy errors (see Figure 7.2-5). However, larger changes in the sensor spectral resolution do have an impact on the overall accuracy. For example, the HIRS sensor used for the TOVS ozone retrievals has wide-bandpass filters, and the total column ozone retrieval accuracy is about 25 DU. Conversely, the EOS-TES instrument has very high spectral resolution, and a projected accuracy of about 1.5 DU (P.D. Brown, Private Communication).

Once ITT was selected as the CrIS sensor contractor it became unlikely for significant changes in the CrIS sensor design from its downselect configuration. While it is possible that slight changes to the spectral resolution may occur, these are not anticipated to have a major impact on the overall OMPS-IR performance.

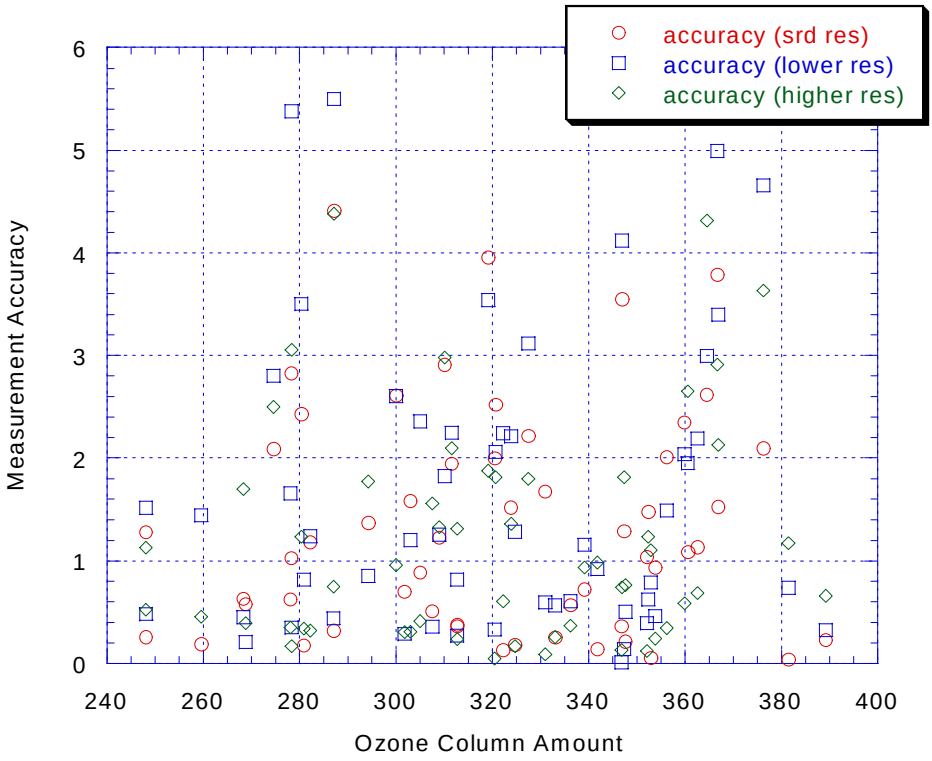


Figure 7.2-5. Impact of spectral resolution on the retrieval accuracy. The three cases shown give accuracy of retrievals for the CrIS SRD spectral resolution, a lower resolution sensor, and a higher resolution sensor. Units for both axes are Dobson Units.

7.2.2.1.2 Sensor Accuracy: Radiometric Calibration Errors

The impact of sensor radiometric calibration errors is an important component in the retrieval accuracy error budget. Despite efforts to thoroughly calibrate the radiances (SDRs) derived from the raw data records (RDRs), the possibility exists that there is a slight bias between the measured radiances and those computed by the forward radiative transfer model. For most infrared sensors this calibration procedure is robust, and an upper limit of 1% of the noise level may be assumed for the calibration errors.

Simulations of calibration errors have been conducted by adding a bias to the simulated sensor data records (SDRs). This bias is computed individually for each CrIS band by determining the mean noise level for that band, multiplying this by the bias factor (e.g. 0.01 for a 1% bias), and adding the result to the simulated radiance. Figure 7.2-6 shows the impact of a +1% bias error to the retrieval accuracy; the impact to the precision is shown in Figure 7.2-7. (Note that a +1% error gives very similar statistics as a –1% error due to the method for calculating accuracy and precision). As was shown in Section 7.3, the choice of covariance has an impact on the retrieval statistics, and a global covariance does a very poor job for the Polar ozone-hole cases (i.e., column ozone amounts less than 210 DU). Overall, however, the addition of sensor bias errors

does not result in significant degradation in the retrieval statistics. Further tests should be conducted once these errors have been adequately characterized by the CrIS contractor.

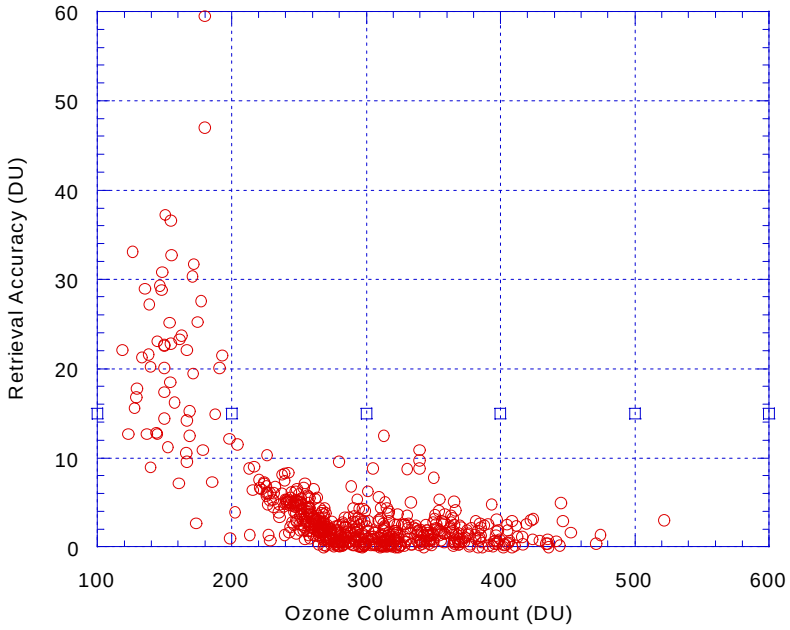


Figure 7.2-6. Impact of a 1% sensor radiometric bias on the retrieval accuracy. The ozone column amounts below 210 DU represent profiles from the enhanced Polar data set. The squares indicate the SRD threshold accuracy requirement (15 DU).

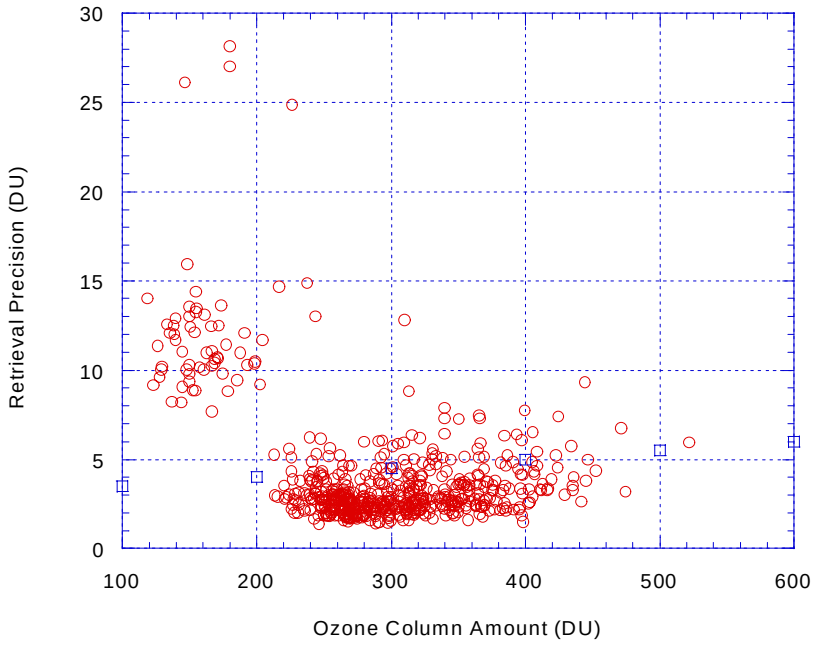


Figure 7.2-7. Impact of sensor radiometric bias on the retrieval precision. The ozone column amounts below 210 DU represent profiles from the enhanced Polar data set. The squares indicate the SRD threshold precision requirement (3 DU + 0.5%).

7.2.2.2 Sensor Impacts to Ozone Retrieval Precision

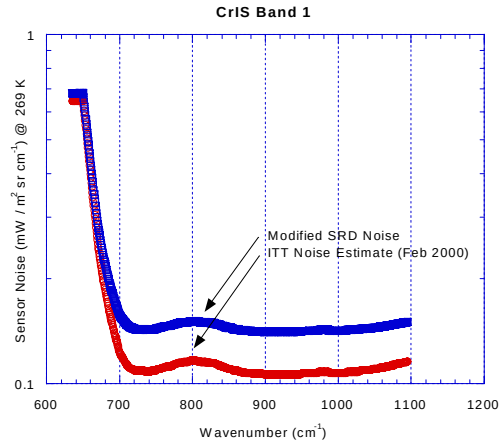
The largest impact of the sensor on the ozone retrieval precision is due to sensor radiometric noise. (The method of including the sensor noise in the simulation and retrieval modules is described in Section 16). Initial tests were done using the parameters given in the CrIS SRD. These tests indicated that there might be a problem in meeting the EDR threshold requirements. Now that the CrIS design has been selected and more information is available about the sensor noise characteristics, the OMPS-IR simulations have been conducted using up-to-date values for the sensor noise. Current results are shown in the baseline retrieval statistics given in Section 7.3.

In addition to the baseline statistics, a test was performed to determine the impact of increasing the noise level from the ITT baseline. At the January 2001 meeting of the CrIS operational algorithm team (the "Sounding OAT" or "SOAT") there was some disagreement about the magnitude and spectral characteristics of the CrIS noise. In particular, in dispute was whether or not the ITT projections of the actual CrIS noise are overly optimistic. Consequently we created a sensor noise file which combines the characteristics of the ITT noise estimates with the nominal noise given in the CrIS SRD prior to the CrIS sensor downselect. These two type of noise, "ITT noise" and "modified SRD noise", are illustrated in Figure 7.2-8. Retrieval simulations were performed with each of these noise values in an attempt to quantify the impact of changes in the sensor noise values. These results, presented in Table 7.2-1, show that the retrieval accuracy is

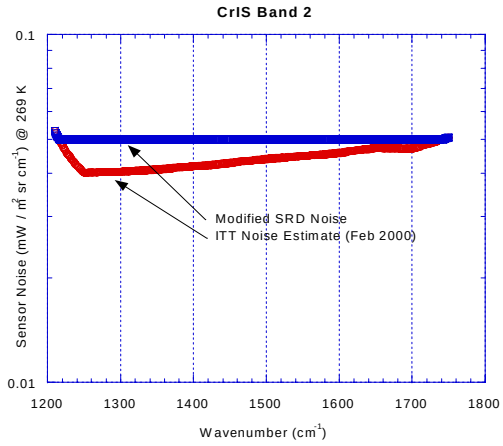
well below the 15 DU threshold requirement, and that the main impact of changing the noise is on the retrieval precision. Noise values larger than those predicted with the "modified SRD" noise values would pose further problems for the ozone retrieval, but are unlikely due to the adverse impacts they would have on the CrIS EDR retrievals.

Table 7.2-1. Accuracy and precision estimates for two different sensor noise values, partitioned by total column ozone amount. (See text for explanation of noise types).

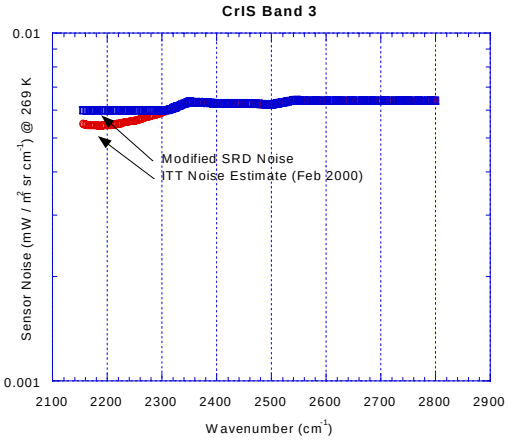
Column O_3 Amount (DU)	ITT Noise		Modified SRD Noise		Threshold Requirement	
	Accuracy (DU)	Precision (DU)	Accuracy (DU)	Precision (DU)	Accuracy (DU)	Precision (DU)
225 ± 25	2.77	2.20	2.91	2.50	15	4.13
275 ± 25	1.36	2.34	1.48	2.68	15	4.38
325 ± 25	1.58	3.22	1.73	3.55	15	4.63
375 ± 25	2.05	3.60	2.20	3.91	15	4.88
425 ± 25	4.16	4.07	4.71	4.47	15	5.13



(a)



(b)



(c)

Figure 7.2-8. Comparisons of ITT noise estimates with a modified version of the CrIS SRD noise values for (a) CrIS band 1, (b) band 2, and (c) band 3.

7.3 Baseline Retrieval Performance Estimates

In this section we present our baseline estimates for the IR total column ozone EDR accuracy and precision. These calculations are based upon a subset of the profiles described in Section 7.1.2 and utilize the latest information about the CrIS sensor and the current version of the OMPS-IR retrieval algorithm.

As described in Section 3.2, the baseline version of the OMPS-IR retrieval algorithm uses CrIS EDRs as input to the OMPS-IR ozone retrieval algorithm. The CrIS SDRs and EDRs are simulated using code obtained from the CrIS team (version 1.2.2, December 2000). Sensor radiometric noise values are given by the ITT baseline of February 2000, which are illustrated in Figure 7.2-8.

The profile data is partitioned into three sets in order to properly calculate covariance matrices and to provide an independent set for testing purposes. The first partition is used to generate the covariance matrices for use by the CrIS EDR code (the "dependent set"). The second set is used by the CrIS code to generate error covariance statistics for the ozone retrieval. The third set is used for the performance simulations (the "independent set"). These three partitions are shown in Figure 7.3-1 (note that the geographical distribution of the dependent set was given in Figure 7.2-4).

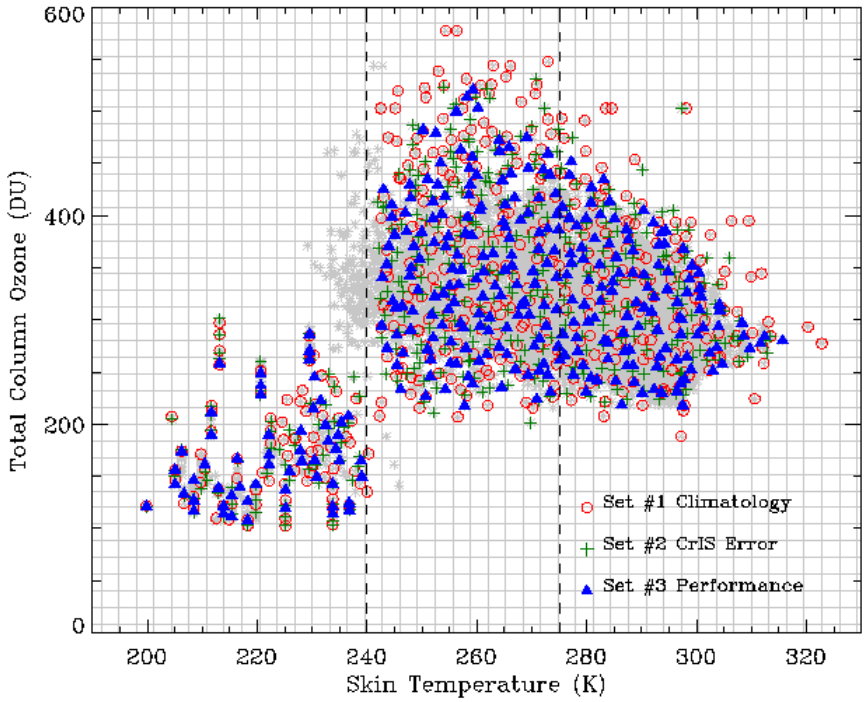


Figure 7.3-1. Selection of profiles used for algorithm testing. Set #1 (red circle) is used to compute the climatological covariance, set #2 (green cross) is used to compute the CrIS error covariance matrix for the ozone retrieval, and set #3 (blue triangle) is for the calculation of the performance statistics.

Performance statistics are computed separately for clear-sky and cloudy-sky retrieval simulations, and are representative of land and ocean conditions for a nadir-viewing sensor. Simulations were also conducted for a 50 degree cross-track scan angle, representing the maximum off-nadir view of the CrIS sensor. The performance results (retrieval accuracy and precision) are calculated using the methods described in Section 7.1.1 and are binned by total column ozone amount as given in Table 7.3-1. A total of 317 different profiles are simulated. Errors due to spectroscopy and sensor radiometric bias are not included in these calculations.

Table 7.3-1. Bins used for retrieval performance statistics.

Total Column Ozone Bin	Range of Ozone (x) in Bin
125 DU	x ≤150 DU
175 DU	150 DU < x ≤ 200 DU
225 DU	200 DU < x ≤ 250 DU
275 DU	250 DU < x ≤ 300 DU
325 DU	300 DU < x ≤ 350 DU
375 DU	350 DU < x ≤ 400 DU
425 DU	x > 400 DU

The current clear-sky, nadir performance results are given in Table 7.3-2 and illustrated graphically in Figure 7.3-2. The baseline retrievals meet the NPOESS accuracy requirement of 15 DU, and for most cases also meet the precision requirement. The retrievals fail to meet the precision requirement for cases with low ozone amounts, particularly those under ozone hole conditions.

Table 7.3-2. Clear-sky, nadir viewing performance statistics. The threshold accuracy requirement is 15 DU. The threshold precision requirement corresponds to 3 DU + 0.5%.

Total Column Ozone	Number of Profiles in Set	Retrieval Accuracy (DU)	Retrieval Precision (DU)	*Threshold Precision (DU)*
125 DU	30	8.35	14.70	*3.6*
175 DU	16	8.04	12.39	*3.9*
225 DU	38	6.72	6.18	*4.1*
275 DU	70	4.56	4.23	*4.4*
325 DU	67	6.72	3.93	*4.6*
375 DU	53	7.98	4.29	*4.9*
425 DU	43	9.08	5.12	*5.1*

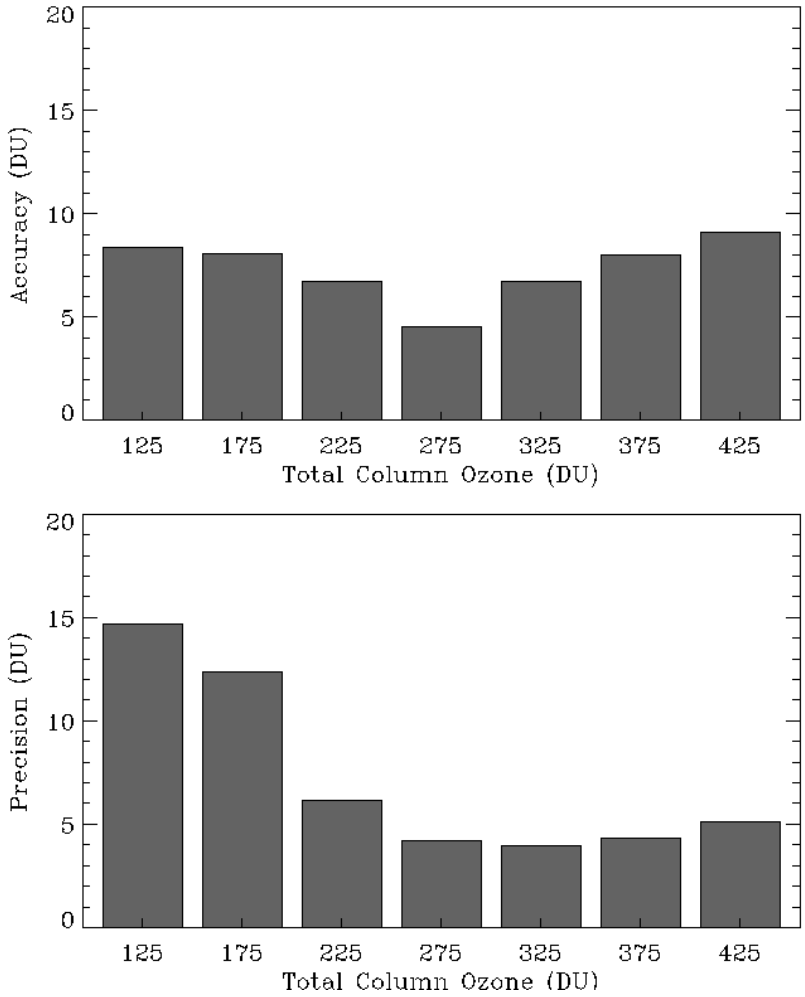


Figure 7.3-2. Clear-sky, nadir viewing performance statistics.

The accuracy and precision for the individual profiles used in the calculation of statistics is given in Figure 7.3-3 as a function of total column ozone amount. The cases below 200 DU clearly exhibit a different type of scatter for the precision than those above 200 DU (the largest precision errors occur for a small group of profiles above 200 DU, but these are offset by the number of cases with very good precision). This systematic difference in the nature of the retrieved amounts can also be seen in Figure 7.3-4, a plot of the retrieval errors as a function of skin temperature and column ozone amount. From this plot it is clear that the profiles near ~250 DU with large errors are actually cold Polar profiles with a relatively large amount of ozone. This may be causing some problems due to lack of representation in the covariance matrix.

An effort is underway to develop a quality control module for assessing the performance of a given retrieval. The goal is to provide a mechanism for pre-retrieval classification and post-retrieval identification of those FOVs that might have inaccurate retrievals. These efforts are discussed in detail in Section 12. Examination of the profiles shown in Figure 7.3-4 indicates that a number of the most inaccurate retrievals could be easily screened, thus improving both the accuracy and precision of the algorithm performance. It is also important to note that the CrIS

FOV is much smaller than the OMPS reporting interval and it is possible that multiple FOVs may be co-added in order to decrease the overall retrieval precision.

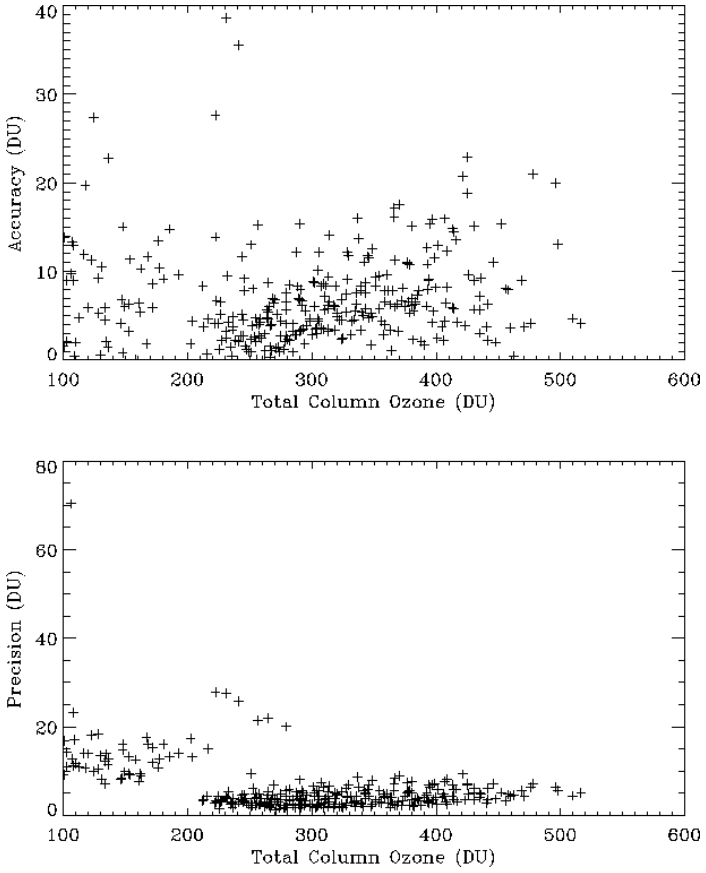


Figure 7.3-3. Accuracy and precision values for the individual profiles comprising the overall performance statistics.

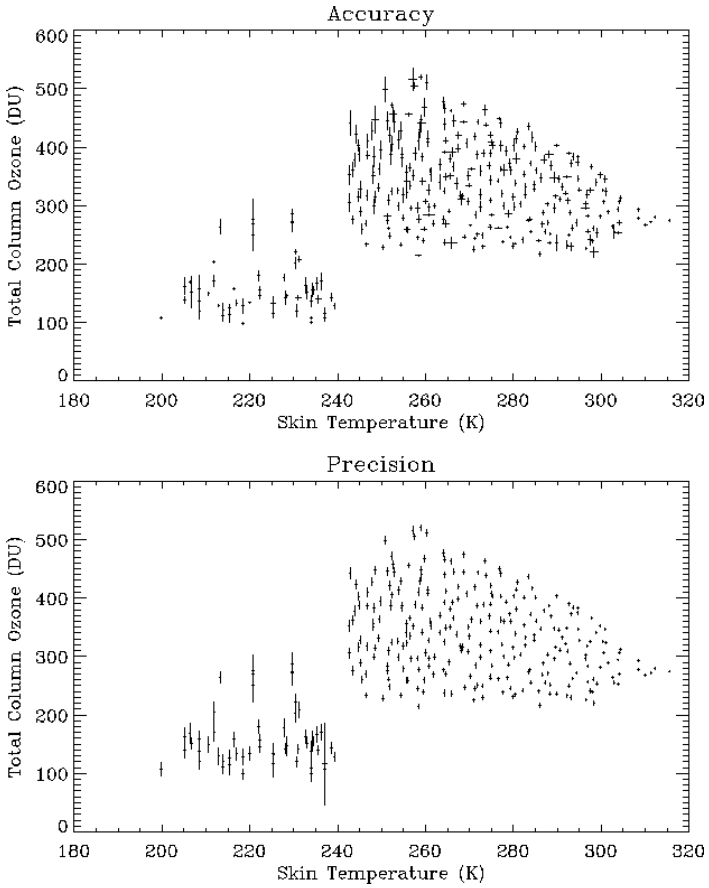


Figure 7.3-4. Accuracy and precision calculations for individual profiles plotted as a function of column ozone and skin temperature. The vertical lines indicate the range of error in column ozone, the horizontal lines indicate the range of error in skin temperature.

Retrieval simulations are also performed for the edge-of-scan condition. For the CrIS sensor this corresponds to a sensor scan angle of 50 degrees. These results are shown in Figure 7.3-5 and Table 7.3-3. For these cases the accuracy and precision results are better than the nadir cases due to increased path length through the atmosphere. However, it is expected that the introduction of clouds will have a larger impact on the edge-of-scan performance than on the nadir retrievals.

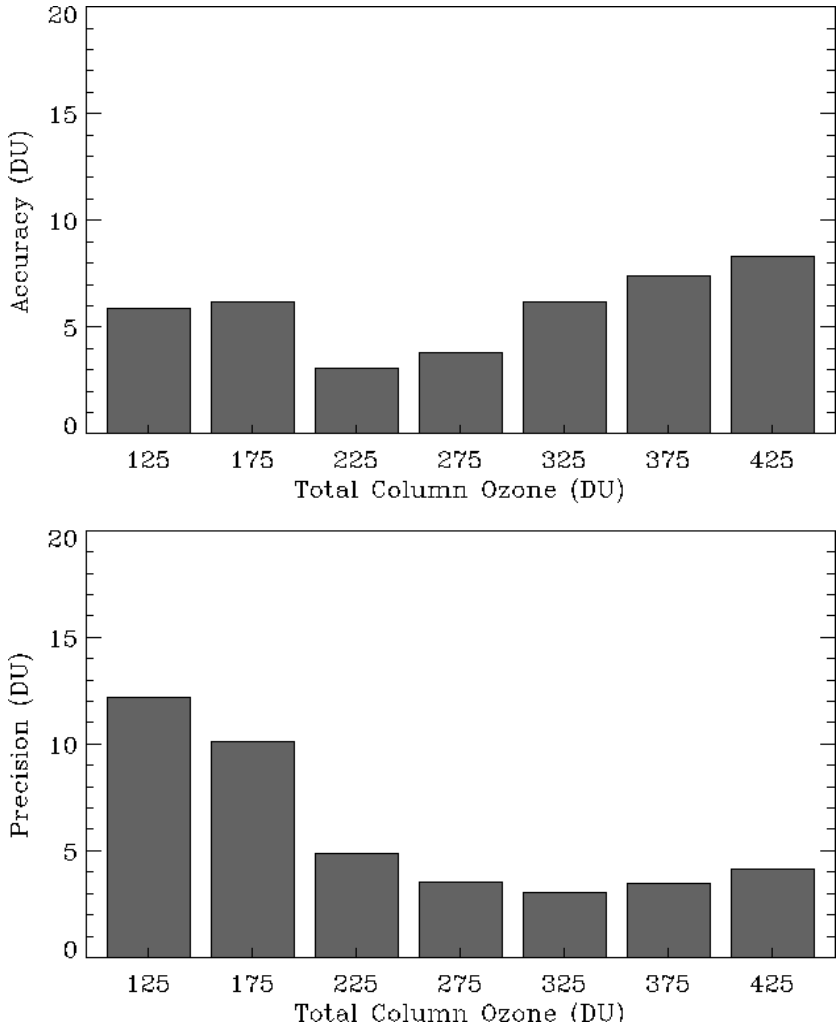


Figure 7.3-5. Accuracy and precision of the IR total column ozone retrieval for the clear sky, edge-of-scan condition.

Table 7.3-3. Clear-sky, edge-of-scan viewing performance statistics. The threshold accuracy requirement is 15 DU. The threshold precision requirement corresponds to 3 DU + 0.5%.

Total Column Ozone	Number of Profiles in Set	Retrieval Accuracy (DU)	Retrieval Precision (DU)	*Threshold Precision (DU)*
125 DU	30	5.89	12.21	*3.6*
175 DU	16	6.17	10.14	*3.9*
225 DU	38	3.06	4.88	*4.1*
275 DU	70	3.79	3.50	*4.4*
325 DU	67	6.17	3.06	*4.6*
375 DU	53	7.40	3.47	*4.9*
425 DU	43	8.32	4.16	*5.1*

8 Test, Calibration, and Validation

Simulated measurements represent the cornerstone of the algorithm development and validation process. Consequently, the development of atmospherically complete and globally representative databases for pre-launch and post-launch testing of the IR-TCA, calibration of the SDRs, and validation of the retrievals is paramount to the success of the IR ozone retrievals. The final CrIS design, in particular the sensor spatial resolution, is key to identifying the databases required for testing. For example, the degree to which cloud and surface information must be specified is sensitive to the spatial resolution selected. Therefore, creation of the test databases is an ongoing process and will continue through the duration of the program.

Aircraft-based radiance measurements in conjunction with a robust determination of the atmospheric state (temperature, pressure, water vapor, ozone) and surface parameters represent a means of algorithm testing under true atmospheric conditions. While these types of measurements do not explicitly match the NPOESS satellite configuration, they are valuable for determining the accuracy and limitations of the forward radiance model.

The final source of validation data is from other satellite sensors. Application of the OMPS-IR algorithm to data from other satellite sensors (e.g. AIRS) will provide a mechanism for validating the operational aspects of the algorithm.

The following sections provide an overview of the types of analysis which should be performed in order to test and validate the OMPS-IR algorithm.

8.1 Overview of Plans

Initial evaluation of the IR total ozone column algorithm utilized the sensor parameters from the CrIS SRD. The tests included the impact of spectral resolution, sensor radiometric noise, sensor calibration (bias) errors, and errors in the IR ozone spectroscopy. Now that the CrIS sensor contractor has been selected, ongoing simulation and test activities use the actual CrIS sensor parameters. Ultimately the EDR algorithms should be coupled to a detailed sensor model to perform complete end-to-end testing of the sensor/algorithm system.

Detailed profile databases, both for testing the algorithm prior to launch and post-launch calibration of the sensor and validation of the algorithm, must reflect both the temporal and spatial variability of ozone. Accurate radiative transfer calculations also require temperature profiles that extend through the stratosphere, and moisture profiles at least to the tropopause. The validity of the algorithm testing also relies on a realistic simulation of clouds and surface parameters. Furthermore, this ensemble of data must accurately reflect the scene variability with latitude, topography and season (see Section 7.1.2). Total column ozone information can be readily obtained from TOMS data and the Dobson spectrophotometer network. However, even at the CrIS spectral resolution the radiances in the 9.6μm ozone band are sensitive to the ozone profile shape (see Section 2.1). Ozone profile measurements have been gathered from SAGE-2 measurements and ozonesondes and used to provide realistic profile shapes for the total column dataset. Because the SAGE-2 information does not extend to sufficiently high altitudes for accurate IR radiative transfer calculations, calculations from a 2-D chemical transport model are used to help fill in the gaps (Kotamarthi et al., 1994; Weisenstein et al., 1996, 1998). Tempera-

ture and moisture profiles come from rocketsondes, radiosondes, satellite-retrieved profiles and numerical model analysis output, all spatially coincident with the ozone measurements used to build the profile database. Representative values of cloud parameters (cloud top height, thickness and optical depth) may be specified from co-located satellite retrievals/measurements and ISCCP databases. Surface parameters, such as skin temperature and surface emissivity, will be obtained from numerical models, surface databases, and satellite measurements. The OMPS-IR test database is described in detail in the AIPT memos AER-AIPT-002 and AER-AIPT-004.

The driving requirement is that databases developed for the testing of the algorithm must reflect the complete variability of factors that contribute to the radiance measured by the CrIS sensor. These data do not need to be coincident in time or space for this stage of testing, although the data do need to be atmospherically realistic, e.g. polar ozone profiles should be linked to polar temperature and moisture profiles. However, the data for the post-launch calibration and validation phases will need to be coincident in time and space to the CrIS overpass.

The profile database is also used for the generation of the atmospheric covariance matrix used by the IR-TCA. The use of enhanced databases in the covariance generation will lead to a more robust performance of the algorithm.

Another task of the algorithm development and testing phase is to study the effect of the first-guess on retrieval accuracy. A first-guess that is close to "truth" does not necessarily lead to the best retrieval of the truth since the retrieval technique may not have enough numerical "momentum" to change the profile during iterations since the differences are so small. However, an unrealistic first guess may lead to non-convergence, or, at the least, additional iterations that waste valuable processing time.

8.2 Simulation and Test Procedures

Using the detailed database described in Section 8.1, the forward RT model is used to generate a data stream that simulates CrIS SDRs. These are used by the IR-TCA to retrieve total column ozone. The "true" values are then compared to the retrieved values to determine the robustness and accuracy of the algorithm.

8.3 Tests using End-to-End Model

The end-to-end model shall consist of a sensor simulation module coupled with the IR-TCA. The end-to-end modeling will be developed in close coordination with the CrIS contractor. These simulations will be more complete and thorough than in Section 8.2 because they will include sensor characterization parameters in a less idealized situation (e.g., the inclusion of off-axis effects in the determination of the sensor spectral response). Using radiances calculated from the detailed atmospheric profile database developed for the simulation studies, the sensor model will generate an SDR data stream that will be used by the IR-TCA to generate the total column ozone EDRs. Comparison of the EDRs with the "truth" data will allow for a determination of the robustness and accuracy of the retrieval algorithm.

8.4 Calibration Data Tests

The key to the post-launch calibration of the IR-TCA will be the construction of a database with spatially and temporally co-located CrIS SDRs and either ground-based or space-based ozone

measurements, temperature and moisture profiles, cloud conditions, and surface parameters. The atmospheric values will be used in the forward RT model, and the results will be compared to the measured CrIS brightness temperatures. In addition, the examination of the average brightness temperatures with respect to scan angle will test for biases. These tests will also be helpful in assessing the calibration and operation of the CrIS sensor.

8.5 Validation through Analysis

Post-launch validation of the ozone retrievals will require the construction of a database with co-located CrIS SDRs and either ground-based or space-based total column ozone measurements. The OMPS-Nadir sensor retrievals will be a valuable source for cross-validation and consistency checks since the pixels will be co-located with the OMPS-IR. The retrieved total column ozone will be compared to the measured values. The validation work will be enhanced including co-located temperature and moisture profiles, cloud conditions, and surface parameters using data from the NPOESS suite of sensors. This will allow retrieval errors to be studied and specific problems identified and corrected. As mentioned in Section 7, a goal of the detailed validation process will be to decrease the ozone band strength uncertainty to at least 2%. The NPP flight of CrIS will provide an excellent validation opportunity.

9 References

9.1 Relevant AER/OMPS Technical Reports

AER-AIPT-001: Trace Gas Impacts on Radiance
AER-AIPT-002: Profile Database Requirements
AER-AIPT-004: Profile Database Implementation
AER-AIPT-006: Solar Geometry
AER-AIPT-007: 2-Stage Algorithm Testing
AER-AIPT-008: Covariance Matrix Tests
AER-AIPT-009: CrIS EDR Code Output Requirements
AER-AIPT-010: Modifications to CrIS EDR Code

9.2 Published References

CrIS ATBD, 2000

Anderson, G.P., S.A. Clough, F.X. Kneizys, J.H. Chetwynd, and E.P. Shettle, AFGL atmospheric constituent profiles (0-120 km), Tech. Rep. AFGL-TR-86-0110, Phillips Lab., Hanscom Air Force Base, Massachusetts, 1986.

Backus, G.E., and J.F. Gilbert, Uniqueness in the inversion of inaccurate gross Earth data, *Philos. Trans. R. Soc. London A*, *266*, 123-192, 1970.

Chahine, M.T., Remote Sounding Cloudy Atmospheres 1: The Single Cloud Layer, *J. Atmos. Science*, 31, 233-243, 1974.

Chahine, M.T., Remote Sounding Cloudy Atmospheres 1: Multiple Cloud Formations, *J. Atmos. Science*, 34, 744-757, 1977.

Chedin, A., N.A. Scott, C. Wahiche, and P. Moulinier, 1985, The improved initialized inversion method: a high resolution physical method for temperature retrievals from TIROS-N series, *J. Clim. Appl. Meteor*., 24, p. 124-143.

Clough, S.A., C. P. Rinsland, and P.D. Brown, Retrieval of tropospheric ozone from simulations of nadir spectral radiances as observed from space, *J. Geophys. Res.*, *100*, 16579-16593, 1995.

Clough, S.A., and M.J. Iacono, Line-by-Line calculations of atmospheric fluxes and cooling rates II: Application to carbon dioxide, ozone, methane, nitrous oxide, and the halocarbons, *J. Geophys. Res.*, *100*, 16519-16535, 1995.

Clough, S.A., M.J. Iacono, and J.-L. Moncet, Line-by-line calculation of atmospheric fluxes and cooling rates: Application to water vapor, *J. Geophys. Res.*, *97*, 15761-15785, 1992.

Clough, S.A., F.X. Kneizys, and R.W. Davies, Line shape and the water vapor continuum, *Atmos. Res.*, *23*, 229-241, 1989.

Curry, J.A., W.B. Rossow, D. Randall, and J.L. Schramm, Overview of Arctic Cloud and Radiation Characteristics, *J. Climate*, 9, 1731-1764, 1996.

Curry, J.A., P.A. Hobbs, M.D. King, D. Randall, P. Minnis, T. Uttal, G.A. Isaac, J. Pinto, et al., FIRE Arctic Clouds Experiment, *Bull. Am. Meteorol. Soc.*, 1999.

Eyre, J.R., Inversion of cloudy satellite sounding radiances by nonlinear optimal estimation. I: Theory and simulation for TOVS, *Q. J. Royal Meteorol. Soc.*, *115*, 1001-1026, 1989.

Gill, E.G., W. Murray, and M.H. Wright, Practical Optimization, Academic, San Diego, Calif., 1981.

Hahn, C.J., S.G. Warren, and J. London, The Effect of Moonlight on Observation of Cloud Cover at Night, and Application to Cloud Climatology, J. Climate, 9, 1429-1446, 1995.

Houghton, J.T., F.W. Taylor, and C.D. Rodgers, Remote Sounding of Atmospheres, Cambridge Planetary Science Series, Cambridge University Press, 1984.

Key, J. and R.G. Barry, Cloud Cover Analysis with Arctic AVHRR Data: 1. Cloud Detection, JGR, 94, 18521-18535, 1989.

Kotamarthi, V.R., M.K.W. Ko, D.K. Weisenstein, J.M. Rodriguez, and N.D. Sze, Effect of Lightning on the concentration of odd nitrogen species in the lower stratosphere: An update, J. Geophys. Res., 99, 8167-8173, 1994.

Kurucz, R.L., Synthetic infrared spectra, in Infrared Solar Physics, edited by D.M. Rabin and J.T. Jefferies, Kluwer, Acad., Norwell, Mass., 1992.

Levenberg, K., A method for the solution of certain problems in least squares, *Q. Appl., Math.*, *2*, 164-168, 1944.

Marks, C.J. and C.D. Rodgers, A retrieval method of atmospheric composition from limb emission measurements, *J. Geophys. Res.*, *98*, 14939-14953, 1993.

Marquardt, D. An algorithm for least squares estimation of nonlinear parameters, *SIAM J. Appl. Math.*, *11*, 431-441, 1963.

McDaniel, A.H., C.A. Cantrell, J.A. Davidson, R.E. Shetter, and J.G. Calvert, The temperature dependent infrared absorption cross sections for the chlorofluorocarbons: CFC-11, CFC-12, CFC-13, CFC-14, CFC-22, CFC-113, CFC-114, and CFC-115, *J. Atmos., Chem.*, *12*, 211-227, 1992.

Miller, A.J., A Review of Satellite Observations of Atmospheric Ozone, *Planet. Space Sci.*, *37*, 1539-1554, 1989.

Miller, S.M., H.E. Snell, and J.-L. Moncet, Simultaneous Retrieval of Middle Atmospheric Temperature and Trace Gas Species Volume Mixing Ratios from Cryogenic Infrared Radiance Instrumentation for Shuttle (CIRRIS 1A), *J. Geophys. Res.*, *104*, 18697-18714, 1999.

Minnis, P., D.R. Doelling, V.Chakrapani, D. Spangenberg, L. Nguyen, R. Palikonda, T. Uttal, R.F. Arduini, and M. Shupe, Cloud coverage during FIRE ACE derived from AVHRR data, *submitted to FIRE ACE Special Issue, J. Geophys. Res.*, November 1999.

Planet, W.G., Crosby, D.S., Lienesch, J.H., and M.L. Hill, Determination of total ozone amount from TIROS radiance measurements, *J. Climate Appl. Meteorol*., *23*, 208-216, 1984.

Rodgers, C.D., Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation, *Rev. Geophys*., *14*, 609-624, 1976.

Rodgers, C.D., Characterization and error analysis of profiles retrieved from remote sounding measurements, *J. Geophys. Res*., *95*, 5587-5595, 1990.

Rossow, W.B., L.C. Garder, and A.A. Lacis, Global, seasonal cloud variations from satellite radiance measurement, I. Sensitivity of analysis, *J. Clim.*, *2*, 419-458, 1989.

Rothman, L.S., C.P. Rinsland, A. Goldman, S.T. Massie, D.P. Edwards, J.-M. Flaud, A. Perrin, C. Camy-Peyret, V. Dana, J.-Y. Mandin, J. Schroeder, A. McCann, R.R. Gamache, R.B. Wattson, K. Yoshino, K.V. Chance, K.W. Jucks, L.R. Brown, V. Nemtchinov, and P. Varanasi, The HITRAN Molecular Spectroscopic Database and HAWKS (HITRAN Atmospheric Workstation): 1996 Edition, *J. Quant. Spect. Radiat. Transfer*, *60*, 665-710, 1998.

Schweiger, A.J., R.W. Lindsay, J.R. Key, and J.A. Francis, Arctic clouds in multiyear satellite data sets, *Geophys. Res. Lett.*, *26*, 1845-1848, 1999.

Smith, W.L., An Improved Method for Calculating Tropospheric Temperature and Moisture from Satellite Radiometer Measurements, *Mon. Weather Rev.*, 96, 387-396, 1968.

Smith, W.L., H.L. Huang, and J.A. Jenny, An Advanced Sounder Cloud Contamination Study, *J. Appl. Meteorology*, 35, 1249-1255, 1996.

Snell, H.E., G.P. Anderson, J. Wang, J.-L. Moncet, J.H. Chetwynd, and S.J. English, Validation of FASE (FASCODE for the Environment) and MODTRAN3: Updates and Comparisons with Clear-Sky Measurements, in Passive Infrared Remote Sensing of Clouds and the Atmosphere III, David K. Lynch and Eric P. Shettle, Editors, Proc. SPIE 2578, 194-204, 1995.

Twomey, S., Introduction to the Mathematics of Inversion in Remote Sensing and Indirect Measurements, *Develop. in Geomath., vol. 3*, Elsevier Sci., New York, 1977.

USGS topograpic map: http://edcwww.cr.usgs.gov/landdaac/gtopo30/gtopo30.html

Weisenstein, D.K., M.K.W. Ko, N.D. Sze and J.M. Rodriguez, Potential impact of SO2 emissions from stratospheric aircraft on ozone, *Geophys. Res. Letters*, *23*, 161-164, 1996.

Weisenstein, D.K., M.K.W. Ko, I.G. Dyominov, G. Pitari, L. Ricciardulli, G. Visconti, and S. Bekki, The effects of sulfur emissions from HSCT aircraft: A 2-D model intercomparison, *J. Geophys. Res.*, *103*, 1527-1547, 1998.

Wylie, D.P., W.P. Menzel, H.M. Woolf, and K.I. Strabala, Four years of global cirrus cloud statistics using HIRS, *J. Climate*, *7*, 1972-1986, 1994.

Wylie, D.P., and W.P. Menzel, Eight years of high cloud statistics using HIRS, *J. Climate*, *12*, 170-184, 1999.

Wilson, L.D., J.A. Curry, and T.P. Ackerman, Satellite Retrieval of Lower-Tropospheric Ice Crystal Clouds in the Polar Regions, *J. Climate*, 6, 1467-1472, 1993.

10 Appendix A: List of Acronyms

ACE	Arctic Clouds Experiment
AER	Atmospheric & Environmental Research, Inc., Lexington, MA
ARM	Atmospheric Radiation Measurement Program (DOE)
ATBD	Algorithm Theoretical Basis Document
CIRRIS	Cryogenic Infrared Radiance Instrument for Shuttle
CKD	Clough-Kneizys-Davies (H_2O Continuum)
CrIS	Cross-track Infrared Sounder
CrIMSS	Cross-track Infrared Microwave Sounder Suite
DOE	Department of Energy
DU	Dobson Units
EDR	Environmental Data Record
EOF	Empirical Orthogonal Function
EOS	Earth Observing System
FASCODE	Fast Atmospheric Signature Code
FIRE	First ISCCP Regional Experiment
FOR	Field-of-Regard
FOV	Field-of-View
HDF	Hierarchical Data Format
HIRS	High Resolution IR Radiation Sounder
HITRAN	High Resolution Transmittance database
HWHM	Half-Width at Half-Maximum
ILS	Instrument Line Shape
IR	Infrared
IR-TCA	IR total column ozone algorithm
ISCCP	International Satellite Cloud Climatology Project
ISO9000	International Organization for Standardization
LBLRTM	Line-by-line radiative transfer model
LOS	Line-of-sight
MSX	Midcourse Space Experiment
NCSA	National Center for Supercomputing Applications
NOAA	National Oceanographic and Atmospheric Administration
NPOESS	National Polar-Orbiting Operational Environmental Satellite System
OAT	Operational Algorithm Team
OE	Optimal Estimation technique
OMPS	Ozone Mapping and Profiler Suite
OPD	Optical Path Difference
OSS	Optical Spectral Sampling algorithm
PDR	Preliminary Design Review
PRCS	Project Revision Control System
RCS	Revision Control System
RDR	Raw Data Record
RT	Radiative Transfer
RTE	Radiative Transfer Equation

SAA	South Atlantic Anomaly
SBUV	Solar Backscatter Ultraviolet
SDR	Sensor Data Record
SGI	Silicon Graphics, Inc.
SHEBA	Surface Heat Budget of the Arctic Ocean
SNR	Signal-to-Noise Ratio
SRD	Sensor Requirements Document
TOA	Top of the Atmosphere
TES	Tropospheric Emission Spectrometer
TIGR	TOVS Initial Guess Retrieval
TOMS	Total Ozone Mapping Spectrometer
TOVS	TIROS Operational Vertical Sounder
TSPR	Total System Performance Responsibility
USGS	United States Geological Survey
vmr	volume mixing ratio

11 Appendix B: Research-Grade Code Development

11.1 Coding Practices

Research grade code has been developed to verify the validity of accuracy and precision claims made for proposed algorithms and to support subsystem (sensor/algorithm) selection/tuning trade studies. These trade studies lead directly to the selection of appropriate algorithms, thresholds, coefficients, and parameters for optimal EDR extraction utilizing the proposed instrument and available external data. Our process model is ISO9001-based, incorporating visibility, configuration management and risk-reduction as key features. These items are critical to ensuring the successful transition of science algorithms to the demanding constraints of an operational system.

Visibility refers to insight into the design and implementation process as well as the proper decomposition of algorithms and process into software modules and the instrumentation of code to support the later transfer of the key technology to an operational environment. In providing key visibility into our code, an operational implementation team can examine our approach and more directly assess the impact of changes to accommodate operational constraints.

Configuration Management refers to change control over methods, code, tests/results and documentation. For each of these items, we have procedures in place to review proposed changes before they are implemented and to evaluate the changes against program objectives.

Risk-reduction is a key theme underlying our development process. We employ standard reviews and evaluations of proposed changes to identify and minimize program risks. Trades are designed and conducted to assess the impact of proposed changes in approach before they are adopted.

11.2 Development Environment

The algorithm development environment utilizes various standards. In particular, our application programs are built in a cross-platform (SUN/SGI) environment, using standard languages and libraries. The source code is in FORTRAN 77/90. The data formats are all self-describing. To support integration of a variety of data sources, we are using ASCII and NetCDF data files. The development environment is characterized by a robust software infrastructure. The workbench for development of prototypes and research grade code, the software infrastructure serves as the breadboard for developing software prototypes. A flexible infrastructure has enabled us to build cross-suite support into our workbench to facilitate post-PDR and TSPR phase trades. Elements of the infrastructure include:

(1) Standard data formats for EDR parameters: We have adopted ASCII as the underlying format for our OMPS-IR algorithm inputs and outputs. While ASCII tends to require more disk space than other formats, this provides easy access to users on all types of platforms. We have also created several NetCDF read routines for accessing the outputs from the CrIS EDR software. NetCDF is a widely used format that enables easy integration and data exchange.

(2) Standard program structure for rapid application development: This policy makes explicit requirements to consider larger subsystem issues (including memory usage and computational efficiency) early in the development cycle.

Developments relate to specific sensor and/or algorithm trades. New work is proposed for peer review. The prime drivers for the evaluation of code modules are impact to EDR performance, specific sensor trades or risk reduction. During the peer review the granularity of the software proposed is evaluated for flexibility of access and visibility into processing. It is presupposed that no substantial code segment exists without instrumentation (key visibility points) which will enable an operational programming group to discover and analyze key processing elements as they are implemented in the science code.

Visibility into methods and calculations is an important aspect of our program development. A standard program structure has been adopted for rapid application development and integration into our baseline of effective alternatives. This policy makes explicit the consideration of software issues early in the development cycle. In general, applications consist of a driver and a library of function/subroutine modules.

The driver program is the backbone of the code modules. While it does not perform any direct calculations, it is responsible for overall program flow and verification of the continued healthy status of the processing. A subroutine is the calculator for one step of an algorithm. It begins by verifying the inputs (checking preconditions and logical assertions). Each subroutine is responsible for verifying the calculation syntax. All control information should be passed explicitly into a subroutine. All subroutines return a result that indicates the status of the operation (success/fail).

Data models are generated with the design of new code to check the requirements against available data. This provides additional visibility into the processing requirements of the algorithms. These abstractions of design should enable an operational team to understand the implementation trade space explored and evaluate other programming alternatives against the required criteria instead of presuming a direct translation is required. This direct support of operational/post-PDR trades is a central theme for continued development of our approach as well as transition to operational constraints.

Configuration management is a key feature for documenting the trade space we examine in the course of the project. At present we use industry-standard UNIX tools (RCS, a Revision Control System and CVS, concurrent version system) for tracking changes to code and executables. In addition to tracking changes to the code, we track changes to the performance of the algorithms. When changes are accepted, a new baseline performance is established. Even if a change is rejected, the results are summarized and archived for future reference. Our build and test environment is a standard SGI/UNIX multi-threaded environment. We use UNIX applications to evaluate the processing performance of our EDR extraction programs.

As previously stated, the code structure consists of a driver and multiple subroutines, each of which is responsible for calculating a partial result. Figure 11.2-1 gives an overview of the algorithm data flow. The code is described in detail in a separate document.

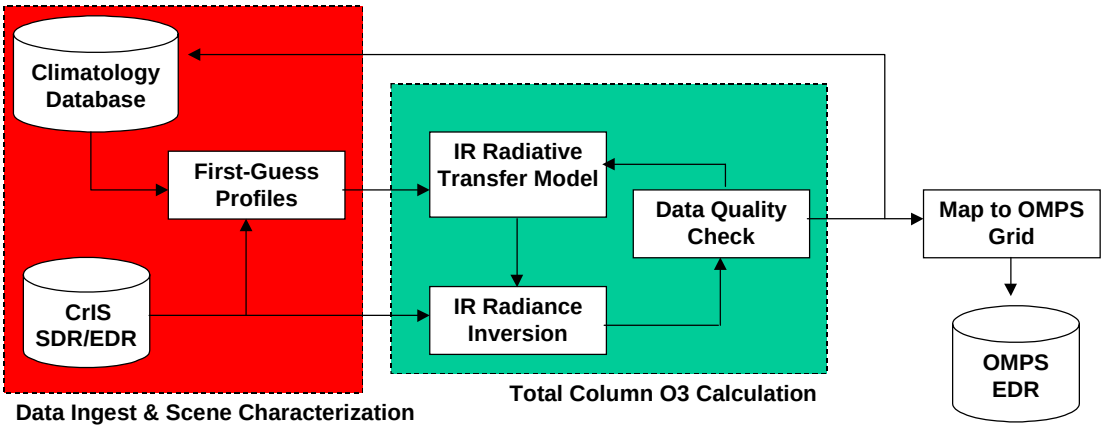


Figure 11.2-1. Algorithm Flow Schematic.

12 Appendix C: Quality Control of the Retrieval Product

Specific details for this section to be completed after RRR3…

Data quality flags will be set for each retrieval, and the output will contain "missing data" flags if no data is available. Using the final retrieved values (ozone, temperature and moisture profiles and surface parameters), the forward RT model calculates the brightness temperatures that should be observed by the sensor. These forward calculated brightness temperatures are compared to the satellite-measured brightness temperatures, and the residual differences are a good check of data quality. Low total residual errors indicate a very consistent retrieval, indicating a reliable retrieval. If the total residual error is high, individual channel residuals indicate where the retrieval is having problems.

In addition to a radiometric means of performing quality control, the error covariance matrix for the retrieval may also be generated.

13 Appendix D: Merging CrIS EDR and OMPS EDR Algorithms

This section describes the practical considerations of merging the algorithm developed by the CrIS team into the algorithms developed for the OMPS IR retrievals in order to compute an accurate assessment of the IR-TCA performance.

An overview of the algorithm structure was provided in Section 3. The IR-TCA uses both SDRs and EDRs as inputs. The CrIS EDR algorithm was provided by the ITT CrIS team in order to accurately simulate the CrIS retrievals for testing the performance of the IR-TCA. The use of the CrIS algorithm for this purpose required several changes to the CrIS code and the OMPS code. The overall set of changes required to couple the OMPS EDR algorithm to the CrIS EDR algorithm are described in detail in AER AIPT memo AER-AIPT-010. Those changes which

impact the outputs from the CrIS EDR code, and thus impose requirements on the CrIS team, are documented in AER AIPT memo AER-AIPT-009.

In summary, ongoing analysis using the December 2000 version of the CrIS EDR algorithm (version 1.2.2) continues to support the benefits of using CrIS EDR products as input into the OMPS-IR total column ozone retrieval. Results from this analysis indicate that to take full advantage of this synergy, two additional derived CrIS products are required by the OMPS-IR algorithm. The first derived product, the CrIS Error Covariance Matrix, is produced by the code for the retrieved parameters (e.g., 20 eigenvectors for temperature and 10 eigenvectors for water vapor) but not currently converted to geophysical profiles or written to an output file. The second derived product, Cloud-Cleared Radiances in the OMPS band (950 – 1095 cm^{-1} region), requires extending the existing CrIS cloud-clearing implementation through the OMPS band and writing the resulting cloud-cleared radiances to an output file.

14 Appendix E: AER Profile Test Data File Format

The data provided for algorithm testing (NOAA-88) has been augmented with Polar profiles to provide a more realistic set of cases for which to test the algorithm (Section 7.1.2). All of these datasets were converted to a 40 level grid for direct use by the OMPS-IR algorithms. This Appendix briefly describes the format of this file.

The OMPS-IR algorithm utilizes a standard 40 level pressure grid (Eyre, 1989) given in Table 11.2-1. The data are stored in a direct access file with a record length of 664 bytes (a total of 116 floating-point elements, each of which are 4 bytes). This format was chosen so that each individual record represents a single profile, and profiles may be selected in any order for a given run. The 166 elements are sufficient to explicitly define the atmospheric and surface conditions for each profile. The order of these elements is given in Table 11.2-2, with the relative humidity flag values listed in Table 11.2-3. Currently the values at which the IR surface and cloud properties are defined (elements 140-146) are: 500, 909, 1111, 2105, 2500, 2857, and 3333 cm^{-1}, and a linear interpolation is performed for the channels between these values.

Table 11.2-1. Pressure grid (40 levels) used by OMPS-IR algorithm.

Reference Pressure Levels (mb):									
0.1	0.2	0.5	1.0	1.5	2.0	3.0	4.0	5.0	7.0
10.0	15.0	20.0	25.0	30.0	50.0	60.0	70.0	85.0	100.0
115.0	135.0	150.0	200.0	250.0	300.0	350.0	400.0	430.0	475.0
500.0	570.0	620.0	670.0	700.0	780.0	850.0	920.0	950.0	1000.0

Table 11.2-2. Variables contained in each record of the profile database.

Element Index	Variable Name	Unit	Definition
1-40	T	K	temperature
41-80	H2O	g/kg	mixing ratio of water vapor
81-120	O3	ppmv	ozone concentration
121	Tskin	K	skin temperature
122	Lat	deg	latitude
123	Lon	deg	longitude
124	Yy		year
125	mm		month
126	dd		day
127	hh		hour
128	mn		minute
129	ss		second
130	Pland		flag (1=land, 0=ocean)
131	Ps	mb	surface pressure
132	epsilon_mw		microwave emissivity
133	Cloud_top_1	mb	type I cloud top pressure
134	Cloud_frac_1		type I cloud fraction
135	Cloud_top_2	mb	type II cloud top pressure
136	Cloud_frac_2		type II cloud fraction
137	col_H2O	kg/m^2	Total column water
138	col_03	DU	Total column ozone
139	RH_flag		RH exceeds 100%
140-146	freq_sfc	wavenumber	IR surface frequency
147-153	emis_sfc		IR surface emissivity
154-160	refl_sfc		IR surface reflectivity
161-162	freq_cld	wavenumber	IR cloud frequency
163-164	emis_cld		IR cloud emissivity
165-166	refl_cld		IR cloud reflectivity

Table 11.2-3. Values for the relative humidity flag (direct access file element 139).

Relative Humidity Flag:
0 - no values in the profile
1 - at least one value at or below 300 mb
2 - at least one value above 300 mb
3 - at least one value above and below 300 mb

15 Appendix F. Algorithm Timing Tests

In addition to requirements for the accuracy and precision of the EDRs, there is a requirement on the overall processing time allowed to convert raw data records (RDRs) into EDRs. The NPOESS requirement for processing time is 20 minutes. However, if this requirement is reduced

to 16 minutes to allow for SDR generation from raw data records (RDRs) and other processes, the time allotted for the retrieval algorithm becomes 0.0047 seconds per FOV (assuming 205,000 FOVs per orbit). While the actual software developed in this phase of the program is not required to execute within the 20 minute processing timeframe, there is a requirement to demonstrate that the algorithms could be converted to operational code that meets that processing time requirements. Several types of tests will be conducted to verify that the algorithm meets this standard.

The research code can be divided into two types of modules, those that will execute with each SDR-to-EDR processing step, and those that execute only once regardless of the number of SDRs to process. An example of the former is the radiative transfer (RT) model, while an example of the latter is the database of coefficients used in the RT model which are read once and stored in memory. In addition there are certain enhancements which can be made to the algorithm to decrease the processing time, such as using the information from an adjacent field-of-view to improve the first-guess and decrease the number of iterations required to compute the EDR.

Two types of tests will be conducted to assess the timing performance of the algorithm. The first test will examine individual modules within the code to determine where most of the processing time occurs. This will be done for a single set of 10 – 20 profiles and will be repeated a sufficient number of times to obtain accurate timing information (i.e. the mean value for a number of retrieval sets in order to average-out the impact of other processes running on the same computer). This test will also allow for an accurate determination of the time required for model initialization (reading of RT databases). The second test will involve the simulation of a single orbit of data and will provide the best estimate of the total processing time. The simulated orbit will contain a realistic sequence of profiles to allow for accurate testing of all aspects of the model timing.

16 Appendix G: Radiometric Noise

The simulation module uses scene-dependant values for the sensor noise. The appropriate NeDN values are provided for each sensor channel at three values of radiometric brightness temperature. The algorithm first converts the noise-free radiance for each sensor channel into the equivalent brightness temperature. A linear interpolation or extrapolation is then performed to convert the channel NeDN values into the value appropriate for the given scene. This number is used with the computer's random number generator to select the appropriate noise value, which is then added to the noise-free radiance.

The retrieval module requires an estimate of the radiometric noise in order to constrain the retrieval solution. This estimate is obtained by converting the "measured" radiance into an equivalent brightness temperature. An interpolation or extrapolation is then used to convert the channel NeDN values into the values appropriate for the scene temperature. This number is used by the algorithm as the sensor noise level (i.e. the S_ε matrix).